

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	28.6	41.4	2360	1	Q65321	Vaccinia virus 7.5	
2	28.6	41.4	2360	1	O85798	Vaccinia virus vec	
3	28.6	41.4	4987	1	O65322	Vaccinia virus pro	
4	28.6	41.4	4987	1	Q86799	Vaccinia virus vec	
5	27.8	40.3	207	1	Q40242	Vaccinia virus ear	
6	26.6	38.6	38	1	Q74458	Poxvirus ATI promo	
7	26.6	38.6	38	1	Q74459	Poxvirus ATI promo	
8	26	37.7	870	1	T63623	DNA encoding L-pro	
9	26	37.7	1081	1	T63622	kPnI-EcoRI fragmen	
10	25.8	37.4	94	1	Q94699	Mycoplasma gallise	
11	25.4	36.8	40	1	Q34776	MlurI-BamHI fragmen	
12	25.4	36.8	40	1	O86231	Recombinant pox vi	
13	25	36.2	40	1	Q68938	Vaccinia virus ear	
14	25	36.2	40	1	Q70568	Vaccinia virus ear	
15	25	36.2	40	1	T48506	sequence of early	
16	25	36.2	70	1	O99648	Synthetic early pr	
17	25	36.2	102	1	Q99624	Swinepox early/lat	
18	25	36.2	108	1	Q99635	Homology vector 52	
19	25	36.2	178	1	Q68946	Homology vector 53	
20	25	36.2	178	1	Q70553	Homology vector 50	
21	25	36.2	178	1	T48503	Homology vector 50	
22	25	36.2	182	1	Q68945	Homology vector 50	
23	25	36.2	182	1	Q70552	Homology vector 50	
24	25	36.2	182	1	T48502	Homology vector 50	
25	25	36.2	205	1	Q99629	Homology vector 53	
26	25	36.2	352	1	Q46676	Sequence Inserted	
27	25	36.2	438	1	Q46679	Sequence Inserted	
28	25	36.2	523	1	Q46677	Sequence Inserted	
29	25	36.2	988	1	T33620	DNA encoding SecA	
30	25	36.2	4177	1	O68943	SfilI fragment cont	
31	25	36.2	4177	1	Q70570	SfilI fragment cont	
32	25	36.2	4177	1	T48510	SfilI fragment cont	
33	24.8	35.9	226	1	T43285	Promoter for infec	

```
PD 14-MAR-1995. 241973.
PF 02-SEP-1993; 241973.
PR 02-SEP-1993; JP-241973.
PA (MITU ) MITSUBISHI KASEI CORP.
DR WPI: 95-144713/19.
PT Antiviral agent comprising component which disrupts viral gene
PT translation - used for the selective inhibition of e.g. Hepatitis
PT C virus, polio:virus and human rhinovirus
PS Disclosure; Page 18; 23pp; Japanese.
CC Q86788 encodes R72800 Hepatitis C virus (HCV) T7N1-19, which
CC disrupts viral gene translation, by preventing the binding of
CC the viral mRNA to the internal ribosome entry site. It was
CC used in the construction of an antiviral agent detecting vector,
CC which comprises a vaccinia virus vector (Q86788), a vector
CC containing T7N1-19 and a firefly luciferase gene. The antiviral
CC agent can be used to treat HCV, polio:virus, cerebral cardio-
CC hepatitis, human rhinovirus and foot and mouth disease viral
CC infections.
SQ Sequence 2360 BP; 858 A; 410 C; 399 G; 693 T;

Query Match 41.4%; Score 28.6; DB 1; Length 2360;
Best Local Similarity 88.6%; Pred. No. 0.25;
Matches 31; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CCAAAATTGAAAACCTAGCTCTATTATTGTCACG 37
Db 1235 CCAAAATTGAAAACCTAGCTCTATTATTGTCACG 1269

RESULT 3
Q65322
ID Q65322 standard; DNA; 4987 BP.
AC Q65322;
DE Vaccinia virus promoter and HCV core protein-luciferase fusion gene.
KW Hepatitis C Virus; HCV; control; Non-A, non-B hepatitis virus;
KW antisense; therapy; inhibition; viral protein precursor;
KW recombinant vaccinia virus; HCV core protein gene; firefly;
KW luciferase reporter gene; fusion construct; ds.
OS Chimeric Vaccinia virus.
OS Chimeric Hepatitis C virus.
OS Chimeric Photinus pyralis.
FH Key Location/Qualifiers
FT cds 1826..4057
FT /tag= a
FT /note= "encodes HCV core protein-luciferase
FT fusion protein"
FT CA2104649-A.
PN 26-FEB-1994.
PD 23-AUG-1993; 104649.
PR 25-AUG-1992; JP-248796.
PR 03-MAR-1993; JP-042736.
PA (SEKI/) SEKI M.
PI Honda Y, Seki M, Yamada E;
DR WPI: 94-154836/19.
DR P-PSDB: R54867.
PT Antisense oligo:nucleotide(s) complementary to the hepatitis C
PT virus genome - are useful as antiviral agents
PS Example 5; Page 243-251; 262pp; English.
CC A recombinant vaccinia virus which codes for a HCV core protein
CC fused to the firefly luciferase enzyme was constructed from PCR
CC amplified fragments. The construct was useful for assaying the
CC inhibitory activity of various antisense oligonucleotides on HCV
CC gene translation.
SQ Sequence 4987 BP; 1520 A; 1052 C; 1083 G; 1332 T;

Query Match 41.4%; Score 28.6; DB 1; Length 4987;
Best Local Similarity 88.6%; Pred. No. 0.31;
Matches 31; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CCAAAATTGAAAACCTAGCTCTATTATTGTCACG 37
Db 1235 CCAAAATTGAAAACCTAGCTCTATTATTGTCACG 1269

RESULT 5
Q40242
ID Q40242 standard; DNA; 207 BP.
AC Q40242;
DE Vaccinia virus early promoter.
DE Vaccinia virus early promoter.
KW HCV; hepatitis C virus; nonB-nonB hepatitis; NANBH; vaccinia virus;
KW VV; cowpox; liver disease; infection; diagnosis; testing; antigen;
KW immunosassay.
OS Synthetic.
PN J05078395-A.
PD 30-MAR-1993.
PR 20-SEP-1991; 241275.
PR 20-SEP-1991; JP-241275.
PA (JAPG ) NIPPON ZEON KK.
DR WPI: 93-140360/17.
DR P-PSDB: R34635.
PT Glyco-co-protein derived from blood disseminating non A non B
PT hepatitis virus - used as diagnostic agent for vaccinia virus
PT infection
PS Disclosure; Page 9; 11pp; Japanese.
CC This sequence is the vaccinia virus early promoter, used when
CC cloning a glycoprotein originating from blood disseminating nonA-nonB
```

CC hepatitis virus (NANBHV) and causing an antigen-antibody reaction with
 CC anti-vaccinia virus. NANBHV originated core and envelope genes were
 CC cloned. The resultant 4.01kb plasmid was named pIK4CE and its sequence
 CC determined. A recombinant vector contg. vaccinia virus (VV) TK gene with
 CC with the VV 7.5kb promoter was prepd. This vector contd. VV originating
 CC TK gene with the synthetic VV early promoter. A second recombinant vector
 CC for the prepn. of vv. having the structural protein gene region of NANBHV
 CC was prepd, and in vv infected cell.
 CC expressed in a VV infected cell.
 SQ Sequence 207 BP; 66 A; 48 C; 38 G; 55 T;

Query Match 40.38; Score 27.8; DB 1; Length 207;
 Best Local Similarity 74.5; Pred. No. 0.25; Mismatches 0; Gaps 0;
 Matches 35; Conservative 0;

QY 5 AAAAATTGAAAACCTAGATCTATTATTCACGCGCGCCATGGCG 51
 |||||
 DB 79 AAAAATTGAAAACCTATCTAATTATTCGAATAGTCGACCATGGCC 125

RESULT 6
 Q74458
 ID Q74458 standard; DNA; 38 BP.
 AC Q74458; 008424.
 DT 22-MAY-1995 (first entry)
 DE Poxvirus ATI promoter, nucleotides -28 to +10.
 KW ATI promoter; poxvirus; pro-phase expression; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 35..38
 FT /label= sticky_end
 FT /tag= a
 FT /note= "see Q74459 for complementary strand"

PN J06237773-A.
 PD 30-AUG-1994.
 PF 21-JAN-1993; 008424.
 PR 22-DEC-1992; JP-342362.
 PA (SHID)/ SHIDA H.
 PA (TOFU) TONEN CORP.
 DR WPI; 94-312814/39.
 PT A poxvirus promoter contg. an ATI promoter and prophase - useful
 PT for the production of exotic genes in large quantities
 PS Claim 2; Page 2; 10pp; Japanese.
 CC Q74458 and Q74459 together form a double-stranded nucleotide
 CC fragment with overhanging sticky ends. The fragment comprises the
 CC ATI promoter and pro-phase expression promoter. Coding sequences
 CC (e.g. heterologous coding sequences or homologous sequences such as
 CC the poxvirus haemagglutinin gene) can be placed downstream of the
 CC promoter fragment for recombinant production of the polypeptides.
 SQ Sequence 38 BP; 15 A; 4 C; 7 G; 12 T;

Query Match 38.68; Score 26.6; DB 1; Length 38;
 Best Local Similarity 87.9; Pred. No. 0.41; Mismatches 0; Gaps 0;
 Matches 29; Conservative 0;

QY 5 AAAAATTGAAAACCTAGATCTATTATTCACG 37
 |||||
 DB 1 AAAAATTGAAAACCTAGTCTAATTATTCACG 33

RESULT 7
 Q74459/C
 ID Q74459 standard; DNA; 38 BP.
 AC Q74459; 008424.
 DT 22-MAY-1995 (first entry)
 DE Poxvirus ATI promoter, nucleotides -32 to +6 (complementary strand).
 KW ATI promoter; poxvirus; pro-phase expression; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 35..38
 FT /tag= a

FT /label= sticky_end
 FT /note= "see Q74458 for complementary strand"
 PN J06237773-A.
 PD 30-AUG-1994.
 PF 21-JAN-1993; 008424.
 PR 22-DEC-1992; JP-342362.
 PA (SHID)/ SHIDA H.
 PA (TOFU) TONEN CORP.
 DR WPI; 94-312814/39.
 PT A poxvirus promoter contg. an ATI promoter and prophase - useful
 PT for the production of exotic genes in large quantities
 PS Claim 2; Page 2; 10pp; Japanese.
 CC Q74458 and Q74459 together form a double-stranded nucleotide
 CC fragment with overhanging sticky ends. The fragment comprises the
 CC ATI promoter and pro-phase expression promoter. Coding sequences
 CC (e.g. heterologous coding sequences or homologous sequences such as
 CC the poxvirus haemagglutinin gene) can be placed downstream of the
 CC promoter fragment for recombinant production of the polypeptides.
 SQ Sequence 38 BP; 12 A; 7 C; 4 G; 15 T;

Query Match 38.68; Score 26.6; DB 1; Length 38;
 Best Local Similarity 87.9; Pred. No. 0.41; Mismatches 0; Gaps 0;
 Matches 29; Conservative 0;

QY 5 AAAAATTGAAAACCTAGATCTATTATTCACG 37
 |||||
 DB 34 AAAAATTGAAAACCTAGTCTAATTATTCACG 2

RESULT 8
 T63623
 ID T63623 standard; DNA; 870 BP.
 AC T63623;
 DT 31-MAY-1997 (first entry)
 DE DNA encoding L-proline-3-hydroxylase.
 KW L-proline-3-hydroxylase; L-proline; cis-3-hydroxy-L-proline;
 KW 2-ketoglutaric acid; ferrous ion; intermediate; drug; food additive; ds.
 OS Streptomyces sp. strain TH1.
 PN WO9627668-A1.
 PD 12-SEP-1996.
 PF 07-MAR-1996; J00558.
 PR 07-MAR-1995; JP-048987.
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 PI Mori H, Ozaki A, Shibasaki T;
 DR WPI; 96-425428/42.
 DR P-PSDB; W09782.
 PT DNA coding for L-proline-3-hydroxylase of microbial origin - for
 PT large scale production of 3-hydroxy-L-proline useful as drug
 PT intermediate or food additive
 PS Claim 3; Page 48; 81pp; Japanese.
 CC The present sequence, derived from pTH30, isolated from Streptomyces
 CC sp., encodes an L-proline-3-hydroxylase (W09782) which converts
 CC L-proline to cis-3-hydroxy-L-proline in the presence of 2-ketoglutaric
 CC acid and ferrous ions. The DNA is used for efficient production of
 CC cis-3-hydroxy-L-proline on an industrial scale for use as an
 CC intermediate in the manufacture of drugs and as a food additive.
 SQ Sequence 870 BP; 173 A; 299 C; 262 G; 136 T;

Query Match 37.78; Score 26; DB 1; Length 870;
 Best Local Similarity 70.0; Pred. No. 1.6; Mismatches 35; Conservative 0; Gaps 0;
 Matches 35; Conservative 0;

QY 18 CTAGATCTATTATTCACGCGCGCCATGGCGCGCCCAACGCG 67
 |||||
 DB 441 CGAGATCTGTTCTCTCGACGCGCGCGCTCCACTCGGCCCTCACTTCG 490

RESULT 9
 T63622
 ID T63622 standard; DNA; 1081 BP.
 AC T63622;

[illegible]

PF 28-FEB-1994; U02252.
PR 26-FEB-1993; US-024156.

us-08-935-377-1.rng

Wed May 31 10:04:50 2000

Job time: 38154 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 11:23:19 ; Search time 621.83 seconds
(without alignments)
14.424 Million cell updates/sec

Title: US-08-935-377-1
Perfect score: 69
Sequence: 1 GCGCAAAATGAAACCTA.....GCCCCGCCCAACGCGCGA 69

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgnl_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgnl_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgnl_6/ptodata/1/ina/5C_COMB.seq.*
4: /cgnl_6/ptodata/1/ina/5D_COMB.seq.*
5: /cgnl_6/ptodata/1/ina/6_COMB.seq.*
6: /cgnl_6/ptodata/1/ina/PCRTUS_COMB.seq.*
7: /cgnl_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	26.4	38.3	40	1	US-07-803-633A-6	Sequence 6, Appl
C 2	26.4	38.3	40	3	US-08-525-742-16	Sequence 16, Appl
C 3	26.4	38.3	96	3	US-08-525-742-30	Sequence 30, Appl
4	26	37.7	870	4	US-08-708-856A-3	Sequence 3, Appl
5	26	37.7	870	5	US-09-287-375-3	Sequence 3, Appl
6	26	37.7	1081	4	US-08-708-856A-13	Sequence 13, Appl
7	26	37.7	1081	5	US-09-287-375-13	Sequence 13, Appl
C 8	25.8	37.4	93	3	US-08-525-742-28	Sequence 28, Appl
9	25.8	37.4	97	3	US-08-525-742-27	Sequence 27, Appl
10	25.4	36.8	40	1	US-07-803-633A-5	Sequence 5, Appl
11	25.4	36.8	40	3	US-08-525-742-15	Sequence 15, Appl
12	25.4	36.8	95	3	US-08-525-742-29	Sequence 29, Appl
13	25	36.2	40	4	US-08-484-575A-8	Sequence 8, Appl
14	25	36.2	40	5	US-08-477-459-8	Sequence 8, Appl
15	25	36.2	40	6	PCT-US94-01826A-8	Sequence 8, Appl
16	25	36.2	40	6	PCT-US94-02252A-8	Sequence 8, Appl
17	25	36.2	70	3	US-08-097-554A-43	Sequence 43, Appl
18	25	36.2	70	5	US-08-480-640A-43	Sequence 43, Appl
19	25	36.2	102	1	US-07-820-154A-10	Sequence 10, Appl
20	25	36.2	102	3	US-08-097-554A-10	Sequence 10, Appl
21	25	36.2	102	5	US-08-480-640A-10	Sequence 10, Appl
22	25	36.2	102	6	PCT-US93-00324-10	Sequence 10, Appl
23	25	36.2	108	1	US-07-820-154A-32	Sequence 32, Appl
24	25	36.2	108	3	US-08-097-554A-32	Sequence 32, Appl
25	25	36.2	108	5	US-08-480-640A-32	Sequence 32, Appl
26	25	36.2	108	6	PCT-US93-00324-32	Sequence 32, Appl

27 25 36.2 111 3 US-08-097-554A-78 Sequence 78, Appl
28 25 36.2 117 5 US-08-480-640A-78 Sequence 78, Appl
29 25 36.2 178 4 US-08-484-575A-16 Sequence 16, Appl
30 25 36.2 178 5 US-08-477-459-16 Sequence 16, Appl
31 25 36.2 178 6 PCT-US94-01826A-16 Sequence 16, Appl
32 25 36.2 178 6 PCT-US94-02252A-16 Sequence 16, Appl
33 25 36.2 182 3 US-08-097-554A-64 Sequence 64, Appl
34 25 36.2 182 4 US-08-484-575A-15 Sequence 15, Appl
35 25 36.2 182 5 US-08-477-459-15 Sequence 15, Appl
36 25 36.2 182 5 US-08-480-640A-64 Sequence 64, Appl
37 25 36.2 182 6 PCT-US94-01826A-15 Sequence 15, Appl
38 25 36.2 182 6 PCT-US94-02252A-15 Sequence 15, Appl
39 25 36.2 185 5 US-08-480-640A-156 Sequence 156, App
40 25 36.2 206 1 US-07-820-154A-21 Sequence 21, Appl
41 25 36.2 206 3 US-08-097-554A-21 Sequence 21, Appl
42 25 36.2 206 5 US-08-480-640A-21 Sequence 21, Appl
43 25 36.2 206 6 PCT-US93-00324-21 Sequence 21, Appl
44 25 36.2 4177 4 US-08-484-575A-12 Sequence 12, Appl
45 25 36.2 4177 5 US-08-477-459-12 Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-07-803-633A-6/c
; Sequence 6, Application US/07803633A
; Patent No. 5369025
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, No. 5369025oru
; APPLICANT: OGAWA, Ryohei
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 No. 5369025th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07803,633A
; FILING DATE: 19911210
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1644-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: Single
; TOPOLOGY: linear
; US-07-803-633A-6

Query Match 38.3%; Score 26.4; DB 1; Length 40;
Best Local Similarity 83.3%; Pred. No. 0.39;
Matches 30; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

5 AAAAAATTGAAAAACTAGACTCTATTATTTCACGCGG 4
39 AAAAAATTGAAAAACTATTCTAATTTATTGCACTCGG 4

39 AAAAAATTGAAAAAACTATTCTTAATTTATTGCACTCGG 4

RESULT 2

US-08-525-742-16/c

; Sequence 16, Application US/08525742

; Patent No. 5871742

; GENERAL INFORMATION:

; APPLICANT: Saito, Shuji

; APPLICANT: Ohkawa, Setsuko

; APPLICANT: Saeki, Sakiko

; APPLICANT: Onsawa, Ikuroh

; APPLICANT: Funato, Hiroko

; APPLICANT: Iritani, Yoshikazu

; APPLICANT: Aoyama, Shigeml

; APPLICANT: Takahashi, Kiyoochito

; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE

; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND

; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL

; TITLE OF INVENTION: AS USE THEREOF

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &

; ADDRESSEE: NAUGHTON

; STREET: 1725 K Street, Suite 1000

; CITY: Washington

; DC

STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525.742
FILING DATE: 25-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-074139
FILING DATE: 31-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-245625
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/00541
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: McLeiland Le-Nrueg
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 950811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-8870357
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

```

US-08-525-742-16
Description: / desc
Synthetic Data
Query Match 38.3%; Score 26.4; DB 3; Length 40;
Best Local Similarity 83.3%; Pred. No. 0.39;
Matches 30; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```



```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; ADDRESSEE: NAUGHTON
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,742
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-074139
; FILING DATE: 31-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-245625
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/00541
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McLealand, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 950811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-659-2930
; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
;
US-08-525-742-28

Query Match 37.4%; Score 25.8; DB 3; Length 93;
Best Local Similarity 81.1%; Pred. No. 0.79;
Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 AAAAATTGAAACTAGATCTATTATTTCACGCGC 41
Db 37 AAAAATTGAAACTATTTCTAATTTATTGCACTGTC 1

RESULT 9
US-08-525-742-27
; Sequence 27, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hiroo
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemi
; APPLICANT: Takahashi, Kiyohito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; TITLE OF INVENTION: AS USE THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
```

```
;
; ADDRESSEE: NAUGHTON
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,742
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-074139
; FILING DATE: 31-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-245625
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/00541
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McLealand, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 950811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-659-2930
; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
;
US-08-525-742-27

Query Match 37.4%; Score 25.8; DB 3; Length 97;
Best Local Similarity 81.1%; Pred. No. 0.8;
Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 AAAAATTGAAACTAGATCTATTATTTCACGCGC 41
Db 61 AAAAATTGAAACTATTTCTAATTTATTGCACTGTC 97

RESULT 10
US-07-803-633A-5
; Sequence 5, Application US/07803633A
; Patent No. 5369025
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, Keyvan
; APPLICANT: LEE, Lucy F.
; APPLICANT: YANAGIDA, No. 5369025oru
; APPLICANT: OGAWA, Ryohel
; APPLICANT: LI, Yi
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR
; TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 301 No. 5369025th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/803,633A
FILING DATE: 19911120
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1644-103P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 241-1300
TELEFAX: (703) 241-2848
TELEX: 248345
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-803-633A-5

Query Match 36.8%; Score 25.4; DB 1; Length 40;
Best Local Similarity 82.9%; Pred. No. 0.88;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 AAAAAATTGAAAACTAGATCTATTATTATTCACGCG 39
|||||
DB 6 AAAAAATTGAAAACTATCTAATTATTATTCACGCG 40
|||||

RESULT 11
US-08-525-742-15
Sequence 15, Application US/08525742
Patent No. 5871742
GENERAL INFORMATION:
APPLICANT: Saito, Shuji
APPLICANT: Ohkawa, Setsuko
APPLICANT: Saeki, Sakiko
APPLICANT: Ohsawa, Ikuroh
APPLICANT: Funato, Hiroo
APPLICANT: Iritani, Yoshikazu
APPLICANT: Aoyama, Shigemi
APPLICANT: Takahashi, Kiyochito
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
ADDRESSEE: NAUGHTON
STREET: 1725 K Street, Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,742
FILING DATE: 25-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-074139
FILING DATE: 31-MAR-1993
PRIOR APPLICATION NUMBER: JP 05-245625
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/803,633A
FILING DATE: 19911120
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1644-103P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 241-1300
TELEFAX: (703) 241-2848
TELEX: 248345
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-803-633A-5

Query Match 36.8%; Score 25.4; DB 1; Length 40;
Best Local Similarity 82.9%; Pred. No. 0.88;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 AAAAAATTGAAAACTAGATCTATTATTATTCACGCG 39
|||||
DB 6 AAAAAATTGAAAACTATCTAATTATTATTCACGCG 40
|||||

RESULT 11
US-08-525-742-15
Sequence 15, Application US/08525742
Patent No. 5871742
GENERAL INFORMATION:
APPLICANT: Saito, Shuji
APPLICANT: Ohkawa, Setsuko
APPLICANT: Saeki, Sakiko
APPLICANT: Ohsawa, Ikuroh
APPLICANT: Funato, Hiroo
APPLICANT: Iritani, Yoshikazu
APPLICANT: Aoyama, Shigemi
APPLICANT: Takahashi, Kiyochito
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
ADDRESSEE: NAUGHTON
STREET: 1725 K Street, Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,742
FILING DATE: 25-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-074139
FILING DATE: 31-MAR-1993
PRIOR APPLICATION NUMBER: JP 05-245625
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP94/00541
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: McLealand, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 950811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-8870357
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "synthetic DNA"
US-08-525-742-29

Query Match 36.8%; Score 25.4; DB 3; Length 95;
Best Local Similarity 82.9%; Pred. No. 1.1;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 5 AAAAATTGAAAACTAGATCTATTATTGCACGC 39
|||||
Db 61 AAAAATTGAAAACTATTCTTAATTATTGCACG 95

RESULT 13
US-08-484-575A-8
Sequence 8, Application US/08484575A
Patent No. 5925358
GENERAL INFORMATION:
APPLICANT: Mark D. Cochran and David E. Junker
TITLE OF INVENTION: Recombinant Fowlpox Viruses and Uses Thereof
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US/08/484,575A
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White Esq, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0450
TELEFAX: (212)391-0525
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-484-575A-8

Query Match 36.2%; Score 25; DB 4; Length 40;
Best Local Similarity 84.8%; Pred. No. 1.2;
Matches 28; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 AAAAATTGAAAACTAGATCTATTATTGCACG 37
|||||
Db 1 AAAAATTGAAAACTATTCTTAATTATTGCACG 33

RESULT 14
US-08-477-459-8
Sequence 8, Application US/08477459
Patent No. 6001369
GENERAL INFORMATION:
APPLICANT: Mark D. Cochran
TITLE OF INVENTION: Recombinant Fowlpox Viruses and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US/08/477,459
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White Esq, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-459-8

Query Match 36.2%; Score 25; DB 5; Length 40;
Best Local Similarity 84.8%; Pred. No. 1.2;
Matches 28; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 AAAAATTGAAAACTAGATCTATTATTGCACG 37
|||||
Db 1 AAAAATTGAAAACTATTCTTAATTATTGCACG 33

RESULT 15
PCT-US94-01826A-8
Sequence 8, Application PC/TUS9401826A
GENERAL INFORMATION:
APPLICANT: Synrio Corporation, et al.
TITLE OF INVENTION: Recombinant Fowlpox Virus S-FPV-043 and Uses
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

us-08-935-377-1.rni

Wed May 31 10:04:50 2000

```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01826A
; FILING DATE: 28-FEB-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; PCT-US94-01826A-8

```

```

Query Match          36.2%; Score 25; DB 6; Length 40;
Best Local Similarity 84.8%; Pred. NO. 1.2;
Matches 28; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy 5 AAAAAATGAAAACTAGATCTATTATTATTCACG 37
    |||||||||||
Db 1 AAAAAATGAAAACTATTCATTAATTATTCACG 33

```

Search completed: May 29, 2000, 22:08:49
Job time: 38730 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 11:21:08 ; Search time 1214.87 Seconds
(without alignments)
432.625 Million cell updates/sec

Title: US-08-935-377-1
Perfect score: 69
Sequence: 1 GCCCAAAATGAAAACTA.....GCCCGCGGCCCAACGGCGGA 69

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 882769 seqs, 3808571567 residues
Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba1.*
2: gb_ba2.*
3: gb_om.*
4: gb_ov.*
5: gb_pat.*
6: gb_ph.*
7: gb_pl1.*
8: gb_pl2.*
9: gb_pr1.*
10: gb_pr2.*
11: gb_pr3.*
12: gb_ro.*
13: gb_sts.*
14: gb_sy.*
15: gb_un.*
16: gb_vi.*
17: em_fun.*
18: em_hum1.*
19: em_hum2.*
20: em_in.*
21: em_om.*
22: em_or.*
23: em_ov.*
24: em_pat.*
25: em_ph.*
26: em_pl.*
27: em_ro.*
28: em_sts.*
29: em_sy.*
30: em_un.*
31: em_vi.*
32: gb_htg1.*
33: gb_htg2.*
34: gb_in1.*
35: gb_in2.*
36: gb_ba1.*
37: gb_ba2.*
38: em_hum3.*
39: em_hum4.*
40: gb_pr4.*
41: gb_htg3.*
42: gb_htg4.*
43: gb_htg5.*

44: gb_htg6.*
45: gb_htg7.*
46: em_htg1.*
47: em_htg2.*
48: em_htg3.*
49: em_hum5.*
50: gb_pl3.*
51: gb_pr5.*
52: gb_htg8.*
53: gb_htg9.*
54: gb_htg10.*
55: gb_htg11.*
56: gb_htg12.*
57: gb_htg13.*
58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	28.6	41.4	2360	5	E08871 A part of g
2	28.6	41.4	4987	5	E08872 DNA encodin
3	27.8	40.3	207	5	E04525 Vaccinia vi
4	26.6	38.5	38	5	E07991 ATI promote
5	26.6	38.5	38	5	E07992 ATI promote
6	26.4	38.3	40	5	AR035285 Sequence
7	26.4	38.3	96	5	AR035298 Sequence
8	26.2	38.0	167879	32	AP000827 Homo sapi
9	26	37.7	2900	1	AB007189 Streptomy
10	26	37.7	186326	52	AC016688 Homo sapi
11	25.8	37.4	93	5	AR035296 Sequence
12	25.8	37.4	97	5	AR035295 Sequence
13	25.4	36.8	36	24	E12061 Synthetic e
14	25.4	36.8	40	5	AR035284 Sequence
15	25.4	36.8	40	5	E08848 Poxvirus ea
16	25.4	36.8	95	5	AR035297 Sequence
17	25.2	36.5	3612	16	EDSORFS
18	25.2	36.5	33213	16	AAVEDSDNA
19	25	36.2	70	5	AR034185 Sequence
20	25	36.2	102	5	AR034169 Sequence
21	25	36.2	108	5	AR034179 Sequence
22	25	36.2	111	5	AR034209 Sequence
23	25	36.2	182	5	AR034200 Sequence
24	25	36.2	206	5	AR034174 Sequence
25	25	36.2	152453	42	AC012130 Homo sapi
26	24.8	35.9	226	24	E12074 Synthetic p
27	24.8	35.9	1433	1	MTU43178 Mycobacteri
28	24.8	35.9	1580	16	AD5004 X02998 Adenovirus
29	24.8	35.9	1580	16	ADRA3 J01969 adenovirus
30	24.8	35.9	3189	5	A52460 Sequence 4
31	24.8	35.9	24171	34	CEC31A11 Z83218 Caenorhabdi
32	24.8	35.9	35935	16	ADRCMPGEN M73260 Mastadenovi
33	24.8	35.9	35937	16	ADRCG J01917 Adenovirus
34	24.8	35.9	69316	55	AC023645 Homo sapi
35	24.8	35.9	157904	56	AC022361 Homo sapi
36	24.8	35.9	217109	53	AC018806 Homo sapi
37	24.6	35.7	24595	1	ATACH5 X00493 Agrobacteri
38	24.6	35.7	24595	5	E00404 T1 plasmid
39	24.6	35.7	24595	5	E00546 E00546 DNA fragmen
40	24.6	35.7	27608	14	AF184978 Binary ve
41	24.6	35.7	72496	44	AC019064 Homo sapi
42	24.6	35.7	172749	40	AC009731 Homo sapi
43	24.4	35.4	86	5	AR035318 Sequence
44	24.4	35.4	91	5	AR035319 Sequence
45	24.4	35.4	2080	2	AFARBCAE M17744 Ralstonia e

ALIGNMENTS

PH	Key	Location/Qualifiers
FH	source	1. .4987
FT		/organism='Artificial sequences' FT CDS
FT	1826. .4057	/product='fusion protein of virus protein and luciferase'.
FT		Location/Qualifiers
FT	1. .4987	/organism='unidentified'
FT		/db_xref="taxon:32644"
BASE COUNT	1511 a 1041 c 1073 g 1314 t	48 others
ORIGIN		
Query Match	41.4%;	Score 28.6; DB 5; Length 4987;
Best Local Similarity	88.6%;	Pred. No. 13;
Matches	31; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
QY	3 CCATAAATTGAAACCTAGATCTATTATTGTCACG 37	
Db	1235 CCATAAATTGAAACCTAGTCTAATTATTGTCACG 1269	
RESULT	3	
E04525	207 bp	DNA PAT 29-SEP-1997
LOCUS		
DEFINITION	Vaccinia virus early promoter.	
ACCESSION	E04525	
VERSION	E04525.1	GI:2172726
KEYWORDS	JP 1993078395-A/6.	
SOURCE	Vaccinia virus.	
ORGANISM	Vaccinia virus	
REFERENCE	1 (bases 1 to 207)	
AUTHORS	Yasuda,K., Sato,T., Nagaya,A. and Kyo,T.	
TITLE	GLUCOPROTEIN, ITS PRODUCTION, RECOMBINED VACCINIA VIRUS FOR ITS	
JOURNAL	PRODUCTION, AND DIAGNOSTIC AGENT USING THE SAME	
COMMENT	Patent: JP 1993078395-A 6 30-MAR-1993;	
	NIPPON ZEON CO LTD	
	OS Vaccinia virus	
	PN JP 1993078395-A/6	
	PD 30-MAR-1993	
	PF 20-SEP-1991 JP 1991241275	
	PI YASUDA KANJI, SATO TAKANORI, NAGAYA ATSUSHI, KYO TSUGUO PC	
	C07K9/00,A61K39/00,A61K39/395,C07K15/14,C12N7/01,C12P21/00, PC	
	G01N33/569,	
	CC GOIN33/576//C12N15/51.(C12P21/00,C12R1:91);	
	CC strandedness: Double;	
	CC topology: Linear;	
FH	Key	Location/Qualifiers
FT	promoter	1. .207
FT		/note='vaccinia virus early promoter'.
FT		Location/Qualifiers
FT	1. .207	/organism="Vaccinia virus"
FT		/db_xref="taxon:10245"
BASE COUNT	66 a 48 c 38 g 55 t	
ORIGIN		
Query Match	40.3%;	Score 27.8; DB 5; Length 207;
Best Local Similarity	74.5%;	Pred. No. 16;
Matches	35; Conservative	0; Mismatches 12; Indels 0; Gaps 0;
QY	5 AAAAATTGAAACCTAGATCTATTATTGTCACGCGCGCCATGGCC 51	
Db	79 AAAAATTGAAACCTATTCTAATTATTGTCACATAGTCGACCATGGCC 125	
RESULT	4	

```

RESULT 1
E08871      2360 bp      RNA      PAT
LOCUS       A part of genomic sequence of Vaccinia virus.
DEFINITION
ACCESSION   E08871.1 GI:2176975
VERSION     JP 1995069899-A/3.
KEYWORDS    Vaccinia virus.
SOURCE      Vaccinia virus.
ORGANISM    Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
REFERENCE   1 (bases 1 to 2360)
AUTHORS     Seki,M., Honda,Y. and Yamada,S.
TITLE       ANTIVIRAL AGENT
JOURNAL     Patent: JP 1995069899-A 3 14-MAR-1995;
COMMENT     MITSUBISHI CHEM CORP
OS          unknown (Vaccinia virus)
PN          JP 1995069899-A/3
PD          14-MAR-1995
PF          02-SEP-1993 JP 1993241973
PI          SEKI MAKOTO, HONDA YOSHIKAZU, YAMADA SUGURU
PC          A61K31/70,A61K48/00,C07K7/00,C12N15/09//C07H21/04; CC
strandedness: Double;
CC topology: Linear;
FH          Key
FH          Location/Qualifiers
FT          1. .2360
FT          /organism='Unclassified'
FT          /clone='PHASE'.

FEATURES             source
BASE COUNT          858 a 410 c 399 g 693 t
ORIGIN

Query Match          41.4%; Score 28.6; DB 5; Length 2360;
Best Local Similarity 88.6%; Pred. No. 12;
Matches 31; Conservative 0; Mismatches 4; Indels 0; Gaps 0

Qy 3 CCAAAATTGAAAACACTAGTCTAATTATTGACG 37
|||||
Db 1235 CCAAAATTGAAAACACTAGTCTAATTATTGACG 1269

RESULT 2
E08872      4987 bp      DNA      PAT
LOCUS       DNA encoding a fusion protein of vaccinia virus protein, HCV protein
DEFINITION
ACCESSION   E08872
VERSION     JP 1995069899-A/4.
KEYWORDS    and Firefly luciferase.
SOURCE      E08872
ORGANISM    E08872.1 GI:2176976
REFERENCE   1 (bases 1 to 4987)
AUTHORS     Seki,M., Honda,Y. and Yamada,S.
TITLE       ANTIVIRAL AGENT
JOURNAL     Patent: JP 1995069899-A 4 14-MAR-1995;
COMMENT     MITSUBISHI CHEM CORP
OS          None
OC          Artificial sequences.
PN          JP 1995069899-A/4
PD          14-MAR-1995
PF          02-SEP-1993 JP 1993241973
PI          SEKI MAKOTO, HONDA YOSHIKAZU, YAMADA SUGURU
PC          A61K31/70,A61K48/00,C07K7/00,C12N15/09//C07H21/04; CC
strandedness: Double;
CC topology: Linear;

```

FT	source	1. .38	/organism='Artificial sequences'.
FT	Location/Qualifiers		
FEATURES	source	1. .38	/organism="unidentified"
		/db_xref="taxon:32644"	
BASE COUNT	12 a	7 c	4 g 15 t
ORIGIN			
Query Match	38.6%;	Score 26.6;	DB 5; Length 38;
Best Local Similarity	87.9%;	Pred. No. 32;	
Matches	29; Conservative	0; Mismatches	4; Indels 0; Gaps 0;
QY	5	AAAAATTGAAAACTAGATCTATTATTGCACG	37
Db	34	AAAAATTGAAAACTAGCTAAATTATTGCACG	2
RESULT	6		
AR035285/c			
LOCUS	AR035285	40 bp	DNA PAT 29-SEP-1999
DEFINITION	Sequence 16 from patent US 5871742.		
ACCESSION	AR035285		
VERSION	AR035285.1	GI:5951953	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 40)		
AUTHORS	Saitoh,S., Ohkawa,S., Saeki,S., Funato,H., Iritani,Y., Aoyama,S. and Takahashi,K.		
TITLE	Recombinant Avipox virus encoding polypeptide of mycoplasma gallisepticum, and utilized a live vaccine		
JOURNAL	Patent: US 5871742-A 16 16-FEB-1999;		
FEATURES	Location/Qualifiers		
source	1. .40		
BASE COUNT	14 a	4 c	5 g 16 t
ORIGIN			
Query Match	38.3%;	Score 26.4;	DB 5; Length 40;
Best Local Similarity	83.3%;	Pred. No. 38;	
Matches	30; Conservative	0; Mismatches	6; Indels 0; Gaps 0;
QY	5	AAAAATTGAAAACTAGATCTATTATTGCACGGG	40
Db	39	AAAAATTGAAAACTATCTAAATTATTGCACTCGG	4
RESULT	7		
AR035298/c			
LOCUS	AR035298	96 bp	DNA PAT 29-SEP-1999
DEFINITION	Sequence 30 from patent US 5871742.		
ACCESSION	AR035298		
VERSION	AR035298.1	GI:5951966	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 96)		
AUTHORS	Saitoh,S., Ohkawa,S., Saeki,S., Funato,H., Iritani,Y., Aoyama,S. and Takahashi,K.		
TITLE	Recombinant Avipox virus encoding polypeptide of mycoplasma gallisepticum, and utilized a live vaccine		
JOURNAL	Patent: US 5871742-A 30 16-FEB-1999;		
FEATURES	Location/Qualifiers		
source	1. .96		
BASE COUNT	44 a	8 c	10 g 34 t
ORIGIN			

wed May 31 10:04:49 2000

Query Match 38.38; Score 26.4; DB 5; Length 96;
Best Local Similarity 83.38; Pred. No. 43;
Matches 30; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 AAAAAATGAAAAATGATCTATTATTATGACGCGG 40
|||||
Db 39 AAAAAATGAAAAATCTATTATTATGACACTCGG 4

RESULT 8
AP000827/c
LOCUS Homo sapiens chromosome 11 clone RP11-716D19 map 11q14, WORKING
DEFINITION DRAFT SEQUENCE, 24 unordered pieces.
ACCESSION AP000827.1 GI:6997662
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens DNA, clone:RP11-716D19.
SOURCE

ORGANISM
(
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 167879)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 167,879 genomic DNA of 11q14
Published Only in Database (1999) in press
2 (bases 1 to 167879)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (03-DEC-1999) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, The Institute of Physical and Chemical Research
(RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1
Kitasato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
Tel:81-42-778-9923, Fax:81-42-778-9924)
On Feb 18, 2000 this sequence version replaced gi:6525269.

----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp

----- Project Information
Center project name: HumDraIt11
Center clone name: RP11-716D19
----- Summary Statistics

Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.950329
Consensus quality: 140861 bases at least Q40
Consensus quality: 149365 bases at least Q30
Consensus quality: 154204 bases at least Q20
Insert size: 156754; sum-of-contigs
Quality coverage: 4.32x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
24 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved

1 14626 contig of 14626 of in length
15127 30176 contig of 15050 of in length
30677 45326 contig of 14650 of in length
45827 58556 contig of 12730 of in length
59057 66715 contig of 7659 of in length
67216 77376 contig of 10161 of in length
77877 88355 contig of 10479 of in length
88856 96474 contig of 7619 of in length
96975 105697 contig of 8723 of in length

14626 contig of 14626 bp in length
15130: gap of 504 bp
15131 30176: contig of 15046 bp in length
30177 30677: gap of 501 bp
30678 45326: contig of 14649 bp in length
45327 45831: gap of 505 bp
45832 58551: contig of 12720 bp in length
58552 59056: gap of 505 bp
59057 66714: contig of 7658 bp in length
66715 67215: gap of 501 bp
67216 77375: contig of 10160 bp in length
77376 77876: gap of 501 bp
77877 88354: contig of 10478 bp in length
88355 88855: gap of 501 bp
88856 96473: gap of 508 bp
96474 106198: gap of 505 bp
105694 106198: gap of 505 bp
106199 114752: contig of 8554 bp in length
114753 115262: gap of 510 bp
115263 122523: contig of 7261 bp in length
122524 123027: gap of 504 bp
123028 130348: contig of 7321 bp in length
130349 130849: gap of 501 bp
130850 137846: contig of 6997 bp in length
137847 138349: gap of 503 bp
138350 142439: contig of 4090 bp in length
142440 142939: gap of 500 bp
142940 146235: contig of 3296 bp in length
146236 146739: gap of 504 bp
146740 150031: contig of 3292 bp in length
150032 150537: gap of 506 bp
150538 153199: contig of 2662 bp in length
153200 153702: gap of 503 bp
153703 155613: contig of 1913 bp in length
155614 156117: gap of 502 bp
156118 157978: contig of 1861 bp in length
157979 158480: gap of 502 bp
158481 160478: contig of 1998 bp in length
160479 160981: gap of 503 bp
160982 162383: contig of 1402 bp in length
162384 162894: gap of 511 bp
162895 164341: contig of 1447 bp in length
164342 164848: gap of 507 bp
164849 166183: contig of 1335 bp in length
166184 166685: gap of 502 bp
166686 167879: contig of 1194 bp in length.

Location/Qualifiers
1..167879
/organism="Homo sapiens"

FEATURES
source

166686

166686

166686

166686

166686

```

CDS
2244..2870
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="BAA22407.1"
/db_xref="GI:2443308"
/translation="MSAEAFDDADIQFAEERAADAVAPRRRREFSTVRCARAL
GELGIPVPPLPPGRHRAPOWPTGYGSMTHCSGYRAAASRLSHSVGDAESAPL
PDGLDLVLGLPERDQVERLGQAQDVFWDRLLFSCKEAYKVFYFPAQRLLGFQDGA
IDIDSGGFFSARFLVPPPLASNGNVPPLTGRWMHRDLILTALPCP"
BASE COUNT      486 a   999 c   945 g   464 t     6 others
ORIGIN

Query Match          37.7%; Score 26; DB 1; Length 2900;
Best Local Similarity 70.0%; Pred. No. 96;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 18 CTAGATCTATTTCACGGCGCCCATGGCCGCCGCACACGCGC 67
| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1752 CGAGATCTGTCTCTCGACGCGCGCGCTCCACTCGCGCGCTCACTTCG 1801

RESULT 10
ACOL16688/LOCUS
DEFINITION Homo sapiens clone RP11-80J14, WORKING DRAFT SEQUENCE, 20 unordered
pieces.
ACCESSION ACOL16688
VERSION ACOL16688.2 GI:6838910
KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 186326)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 186326)
Waterston,R.H.
Direct Submission
Submitted (04-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jan 31, 2000 this sequence version replaced gi:6524391.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0080J14
----- Summary Statistics -----
Sequencing vector: M13; 80%
Chemistry: Dye-primer ET; 80% of reads
Assembly: Dye-terminator Big Dye; 20% of reads
Consensus quality: 174077 bases at least Q40
Consensus quality: 178110 bases at least Q30
Consensus quality: 180386 bases at least Q20
Insert size: 174000; agarose-fp
Insert size: 186326; sum-of-contigs
Quality coverage: 4.31 in Q20 bases; agarose-fp
Quality coverage: 4.03 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```

```

CDS
2244..2870
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="BAA22407.1"
/db_xref="GI:2443308"
/translation="MSAEAFDDADIQFAEERAADAVAPRRRREFSTVRCARAL
GELGIPVPPLPPGRHRAPOWPTGYGSMTHCSGYRAAASRLSHSVGDAESAPL
PDGLDLVLGLPERDQVERLGQAQDVFWDRLLFSCKEAYKVFYFPAQRLLGFQDGA
IDIDSGGFFSARFLVPPPLASNGNVPPLTGRWMHRDLILTALPCP"
BASE COUNT      486 a   999 c   945 g   464 t     6 others
ORIGIN

Query Match          37.7%; Score 26; DB 1; Length 2900;
Best Local Similarity 70.0%; Pred. No. 96;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 18 CTAGATCTATTTCACGGCGGCCCATGGCCCGCCGCACACGCGC 67
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1752 CGAGATCTGTCTCTCGACGCGCGCGCTCCACTCGCGCGCTCACTCG 1801
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
ACOL16688/c
LOCUS
DEFINITION
Homo sapiens clone RP11-80J14, WORKING DRAFT SEQUENCE, 20 unordered
pieces.
ACCESSION ACOL16688
VERSION ACOL16688.2 GI:6838910
KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 186326)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 186326)
Waterston,R.H.
Direct Submission
Submitted (04-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jan 31, 2000 this sequence version replaced gi:6524391.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0080J14
----- Summary Statistics -----
Sequencing vector: M13; 80%
Chemistry: Dye-primer ET; 80% of reads
Assembly: Dye-terminator Big Dye; 20% of reads
Consensus quality: 174077 bases at least Q40
Consensus quality: 178110 bases at least Q30
Consensus quality: 180386 bases at least Q20
Insert size: 174000; agarose-fp
Insert size: 186326; sum-of-contigs
Quality coverage: 4.31 in Q20 bases; agarose-fp
Quality coverage: 4.03 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
```

* as soon as it is available and the accession number will
* be preserved.
1 1456: contig of 1456 bp in length
* gap of unknown length
* 1457 2971: contig of 1515 bp in length
* gap of unknown length
* 2972 5544: contig of 2573 bp in length
* gap of unknown length
* 5545 7248: contig of 1704 bp in length
* gap of unknown length
* 7249 8398: contig of 1150 bp in length
* gap of unknown length
* 8399 10354: contig of 1956 bp in length
* gap of unknown length
* 10355 12998: contig of 2644 bp in length
* gap of unknown length
* 12999 16483: contig of 3485 bp in length
* gap of unknown length
* 16484 20173: contig of 3690 bp in length
* gap of unknown length
* 20174 27635: contig of 7462 bp in length
* gap of unknown length
* 27636 37720: contig of 10085 bp in length
* gap of unknown length
* 37721 46381: contig of 8661 bp in length
* gap of unknown length
* 46382 55849: contig of 9468 bp in length
* gap of unknown length
* 55850 63824: contig of 7975 bp in length
* gap of unknown length
* 63825 76918: contig of 13094 bp in length
* gap of unknown length
* 76919 90722: contig of 13804 bp in length
* gap of unknown length
* 90723 109347: contig of 18625 bp in length
* gap of unknown length
* 109348 126404: contig of 17057 bp in length
* gap of unknown length
* 126405 148648: contig of 22244 bp in length
* gap of unknown length
* 148649 186326: contig of 37678 bp in length.
Location/Qualifiers
1..186326
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Rp11-80J14"
BASE COUNT 54067 a 40241 c 39737 g 52239 t 42 others
ORIGIN
Query Match 37.7%; Score 26; DB 52; Length 186326;
Best Local Similarity 62.1%; Pred. No. 1.7e-02;
Matches 41; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 3 CCAAAATTGAAACTAGATCTATTATTCACGCGCGCATGGCGCGGCCCAA 62
Db 5570 CAAAGATAAAACTCTTCTTTTAAAGCACAGGCGCCCATGCCCCCTCTCCCAA 5511
QY 63 CGCGCG 68
Db 5510 GGCGCG 5505
RESULT 11
LOCUS AR035296/c 93 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 28 from patent US 5871742.
ACCESSION AR035296
VERSION AR035296.1 GI:5951964
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 93)
AUTHORS Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y.,
Aoyama,S. and Takahashi,K.
TITLE Recombinant Avipox virus encoding polypeptide of mycoplasma
gallisepticum, and utilized a live vaccine
JOURNAL Patent: US 5871742-A 28 16-FEB-1999;
FEATURES Location/Qualifiers
source 1..93
BASE COUNT 43 a 7 c 10 g 33 t
ORIGIN
Query Match 37.4%; Score 25.8; DB 5; Length 93;
Best Local Similarity 81.1%; Pred. No. 69;
Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 5 AAAAATTGAAACTAGATCTATTATTCACGCGCGC 41
Db 37 AAAAATTGAAACTATCTCTAATTATTGCACTCGTC 1
RESULT 12
LOCUS AR035295 97 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 27 from patent US 5871742.
ACCESSION AR035295
VERSION AR035295.1 GI:5951963
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 97)
AUTHORS Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y.,
Aoyama,S. and Takahashi,K.
TITLE Recombinant Avipox virus encoding polypeptide of mycoplasma
gallisepticum, and utilized a live vaccine
JOURNAL Patent: US 5871742-A 27 16-FEB-1999;
FEATURES Location/Qualifiers
source 1..97
BASE COUNT 34 a 11 c 8 g 44 t
ORIGIN
Query Match 37.4%; Score 25.8; DB 5; Length 97;
Best Local Similarity 81.1%; Pred. No. 69;
Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 5 AAAAATTGAAACTAGATCTATTATTCACGCGCGC 41
Db 61 AAAAATTGAAACTATCTCTAATTATTGCACTCGTC 97
RESULT 13
E12061 standard; DNA; UNC; 36 BP.
ID E12061
XX E12061;
AC E12061;
XX E12061.1
SV E12061.1
XX 08-OCT-1997 (Rel. 52, Created)
DT 08-OCT-1997 (Rel. 52, Last updated, Version 1)
XX Synthetic early promoter sequence.
DE JP 1996242869-A/2.
XX JP 1996242869-A/2.
XX unidentified
OS unclassified.
OC unclassified.
XX
RN [1]


```
RP 1-36
RA Kamogawa K., Ogawa R., Yamaguchi T., Hirai K.;
RT "RECOMBINANT POXVIRUS AND VACCINE COMPRISING THE SAME";
RL Patent number JP 1996242869-A/2, 24-SEP-1996.
RL NIPPON ZEON CO LTD.
XX
CC OS None
CC OC Artificial sequences.
CC PN JP 1996242869-A/2
CC PD 24-SEP-1996
CC PF 16-JAN-1996 JP 1996023125
CC PR 13-JAN-1995 JP 95P 21249
CC PI KAMOGAWA KOICHI, OGAWA RYOHEI, YAMAGUCHI TAKESHI,
CC HIRAI KATSUYA
CC PC C12N15/09,A61K39/12,C12N7/00;
CC CC strandedness: Double;
CC CC topology: Linear;
CC CC hypothetical: No;
CC CC anti-sense: No;
CC CC key Location/Qualifiers
CC FH
CC FT source 1..36
CC FT /organism="Artificial sequences"
XX
FH Key Location/Qualifiers
FH source 1..36
FH /db_xref="taxon:32644"
FH /organism="unidentified"
XX
SQ Sequence 36 BP; 15 A; 5 C; 3 G; 13 T; 0 other;

Query Match 36.8%; Score 25.4; DB 24; Length 36;
Best Local Similarity 82.9%; Pred. No. 82;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 AAAAATTGAAAACTAGATCTATTATTATTCACGCG 39
|||||
DB 2 AAAAATTGAAAACTATCTTAATTTATTCACGCG 36
|||||

RESULT 14
AR035284
LOCUS AR035284 40 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 15 from patent US 5871742.
ACCESSION AR035284
VERSION AR035284.1 GI:5951952
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 40)
AUTHORS Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y.,
Aoyama,S. and Takahashi,K.
TITLE Recombinant Avipox virus encoding polypeptide of mycoplasma
gallisepticum, and utilized a live vaccine
JOURNAL Patent: US 5871742-A 15 16-FEB-1999;
FEATURES
source 1..40
BASE COUNT 15 a 7 c 5 g 13 t
ORIGIN

Query Match 36.8%; Score 25.4; DB 5; Length 40;
Best Local Similarity 82.9%; Pred. No. 84;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 AAAAATTGAAAACTAGATCTATTATTATTCACGCG 39
|||||
DB 6 AAAAATTGAAAACTATCTTAATTTATTCACGCG 40
|||||

RESULT 15
E08848
LOCUS E08848 40 bp DNA PAT 29-SEP-1997
DEFINITION Poxvirus early promoter.
ACCESSION E08848
VERSION E08848.1 GI:2176952
KEYWORDS JP 199567655-A/1.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Yamaguchi,T., Fukushi,H., Hirai,K., Aoyama,S., Yamaguchi,T.,
Iritani,K., Hayashi,Y., Ogawa,R., Takamura,C. and Kamogawa,K.
TITLE PROMOTER FOR COMBINED POXVIRUS AND RECOMBINANT POXVIRUS HAVING THE
JOURNAL Patent: JP 1995067655-A 1 14-MAR-1995;
NIPPON ZEON CO LTD, SHIONOGI & CO LTD
COMMENT
OS None
OC Artificial sequences.
PN JP 1995067655-A/1
PD 14-MAR-1995
PF 30-AUG-1993 JP 1993238953
PI YAMAGUCHI TAKESHI, FUKUSHI HIDEKI, HIRAI KATSUYA, PI AOYAMA
SHIGEMI,
PI YAMAGUCHI TAKESHI, IIRITANI KOICHI, HAYASHI YUKIHIRO, PI
OGAWA RYOHEI,
PI TAKAMURA CHIZUKO, KAMOGAWA KOICHI
PC C12N15/09,A61K39/275,C12N7/00.(C12N15/09,C12R1:92),(C12N7/00,
PC C12R1:92);
CC strandedness: Double;
CC topology: Linear;
CC key Location/Qualifiers
FH
FH source 1..40
FH /organism="Artificial sequences".
FEATURES
source 1..40
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 15 a 7 c 5 g 13 t
ORIGIN

Query Match 36.8%; Score 25.4; DB 5; Length 40;
Best Local Similarity 82.9%; Pred. No. 84;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 AAAAATTGAAAACTAGATCTATTATTATTCACGCG 39
|||||
DB 6 AAAAATTGAAAACTATCTTAATTTATTCACGCG 40
|||||

Search completed: May 29, 2000, 21:33:53
Job time: 36765 sec
```

us-08-935-377-1.rge

Wed May 31 10:04:49 2000

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 17:05:30 ; Search time 2276.24 Seconds
(without alignments)
135.250 Million cell updates/sec

Title: US-08-935-377-1
Perfect score: 69
Sequence: 1 GCCCAAAATTGAAAACTA.....GCCCGCGCCGCAACGGCGGA 69

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5142629 seqs, 2230885800 residues

Total number of hits satisfying chosen parameters: 10285240

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents.NA.*

- 1: /cgnl_6/ptodata/1/pna/PCTUS_COMB.seq.*
- 2: /cgnl_6/ptodata/1/pna/US06_COMB.seq.*
- 3: /cgnl_6/ptodata/1/pna/US07_COMB.seq.*
- 4: /cgnl_6/ptodata/1/pna/US080_COMB.seq.*
- 5: /cgnl_6/ptodata/1/pna/US081A_COMB.seq.*
- 6: /cgnl_6/ptodata/1/pna/US081B_COMB.seq.*
- 7: /cgnl_6/ptodata/1/pna/US081C_COMB.seq.*
- 8: /cgnl_6/ptodata/1/pna/US081D_COMB.seq.*
- 9: /cgnl_6/ptodata/1/pna/US082A_COMB.seq.*
- 10: /cgnl_6/ptodata/1/pna/US082B_COMB.seq.*
- 11: /cgnl_6/ptodata/1/pna/US082C_COMB.seq.*
- 12: /cgnl_6/ptodata/1/pna/US083A_COMB.seq.*
- 13: /cgnl_6/ptodata/1/pna/US083B_COMB.seq.*
- 14: /cgnl_6/ptodata/1/pna/US084A_COMB.seq.*
- 15: /cgnl_6/ptodata/1/pna/US084B_COMB.seq.*
- 16: /cgnl_6/ptodata/1/pna/US084C_COMB.seq.*
- 17: /cgnl_6/ptodata/1/pna/US085_COMB.seq.*
- 18: /cgnl_6/ptodata/1/pna/US086_COMB.seq.*
- 19: /cgnl_6/ptodata/1/pna/US087A_COMB.seq.*
- 20: /cgnl_6/ptodata/1/pna/US087B_COMB.seq.*
- 21: /cgnl_6/ptodata/1/pna/US087C_COMB.seq.*
- 22: /cgnl_6/ptodata/1/pna/US088A_COMB.seq.*
- 23: /cgnl_6/ptodata/1/pna/US088B_COMB.seq.*
- 24: /cgnl_6/ptodata/1/pna/US088C_COMB.seq.*
- 25: /cgnl_6/ptodata/1/pna/US088D_COMB.seq.*
- 26: /cgnl_6/ptodata/1/pna/US089A_COMB.seq.*
- 27: /cgnl_6/ptodata/1/pna/US089B_COMB.seq.*
- 28: /cgnl_6/ptodata/1/pna/US089C_COMB.seq.*
- 29: /cgnl_6/ptodata/1/pna/US089D_COMB.seq.*
- 30: /cgnl_6/ptodata/1/pna/US089E_COMB.seq.*
- 31: /cgnl_6/ptodata/1/pna/US089F_COMB.seq.*
- 32: /cgnl_6/ptodata/1/pna/US089G_COMB.seq.*
- 33: /cgnl_6/ptodata/1/pna/US089H_COMB.seq.*
- 34: /cgnl_6/ptodata/1/pna/US089I_COMB.seq.*
- 35: /cgnl_6/ptodata/1/pna/US090A_COMB.seq.*
- 36: /cgnl_6/ptodata/1/pna/US090B_COMB.seq.*
- 37: /cgnl_6/ptodata/1/pna/US090C_COMB.seq.*
- 38: /cgnl_6/ptodata/1/pna/US090D_COMB.seq.*
- 39: /cgnl_6/ptodata/1/pna/US090E_COMB.seq.*
- 40: /cgnl_6/ptodata/1/pna/US091A_COMB.seq.*
- 41: /cgnl_6/ptodata/1/pna/US091B_COMB.seq.*
- 42: /cgnl_6/ptodata/1/pna/US091C_COMB.seq.*
- 43: /cgnl_6/ptodata/1/pna/US092A_COMB.seq.*

- 44: /cgnl_6/ptodata/1/pna/US092B_COMB.seq.*
- 45: /cgnl_6/ptodata/1/pna/US092C_COMB.seq.*
- 46: /cgnl_6/ptodata/1/pna/US092D_COMB.seq.*
- 47: /cgnl_6/ptodata/1/pna/US092E_COMB.seq.*
- 48: /cgnl_6/ptodata/1/pna/US093A_COMB.seq.*
- 49: /cgnl_6/ptodata/1/pna/US093B_COMB.seq.*
- 50: /cgnl_6/ptodata/1/pna/US093C_COMB.seq.*
- 51: /cgnl_6/ptodata/1/pna/US093D_COMB.seq.*
- 52: /cgnl_6/ptodata/1/pna/US093E_COMB.seq.*
- 53: /cgnl_6/ptodata/1/pna/US093F_COMB.seq.*
- 54: /cgnl_6/ptodata/1/pna/US094A_COMB.seq.*
- 55: /cgnl_6/ptodata/1/pna/US094B_COMB.seq.*
- 56: /cgnl_6/ptodata/1/pna/US094C_COMB.seq.*
- 57: /cgnl_6/ptodata/1/pna/US095_COMB.seq.*
- 58: /cgnl_6/ptodata/1/pna/US6000_COMB.seq.*
- 59: /cgnl_6/ptodata/1/pna/US6001A_COMB.seq.*
- 60: /cgnl_6/ptodata/1/pna/US6001B_COMB.seq.*
- 61: /cgnl_6/ptodata/1/pna/US6002A_COMB.seq.*
- 62: /cgnl_6/ptodata/1/pna/US6002B_COMB.seq.*
- 63: /cgnl_6/ptodata/1/pna/US6002C_COMB.seq.*
- 64: /cgnl_6/ptodata/1/pna/US6003A_COMB.seq.*
- 65: /cgnl_6/ptodata/1/pna/US6003B_COMB.seq.*
- 66: /cgnl_6/ptodata/1/pna/US6004A_COMB.seq.*
- 67: /cgnl_6/ptodata/1/pna/US6004B_COMB.seq.*
- 68: /cgnl_6/ptodata/1/pna/US6004C_COMB.seq.*
- 69: /cgnl_6/ptodata/1/pna/US6005_COMB.seq.*
- 70: /cgnl_6/ptodata/1/pna/US6006_COMB.seq.*
- 71: /cgnl_6/ptodata/1/pna/US6007_COMB.seq.*
- 72: /cgnl_6/ptodata/1/pna/US6008A_COMB.seq.*
- 73: /cgnl_6/ptodata/1/pna/US6008B_COMB.seq.*
- 74: /cgnl_6/ptodata/1/pna/US6008C_COMB.seq.*
- 75: /cgnl_6/ptodata/1/pna/US6009A_COMB.seq.*
- 76: /cgnl_6/ptodata/1/pna/US6009B_COMB.seq.*
- 77: /cgnl_6/ptodata/1/pna/US6010A_COMB.seq.*
- 78: /cgnl_6/ptodata/1/pna/US6010B_COMB.seq.*
- 79: /cgnl_6/ptodata/1/pna/US6011A_COMB.seq.*
- 80: /cgnl_6/ptodata/1/pna/US6011B_COMB.seq.*
- 81: /cgnl_6/ptodata/1/pna/US6012A_COMB.seq.*
- 82: /cgnl_6/ptodata/1/pna/US6012B_COMB.seq.*
- 83: /cgnl_6/ptodata/1/pna/US6013A_COMB.seq.*
- 84: /cgnl_6/ptodata/1/pna/US6013B_COMB.seq.*
- 85: /cgnl_6/ptodata/1/pna/US6014A_COMB.seq.*
- 86: /cgnl_6/ptodata/1/pna/US6014B_COMB.seq.*
- 87: /cgnl_6/ptodata/1/pna/US6014C_COMB.seq.*
- 88: /cgnl_6/ptodata/1/pna/US6015_COMB.seq.*
- 89: /cgnl_6/ptodata/1/pna/US6016A_COMB.seq.*
- 90: /cgnl_6/ptodata/1/pna/US6016B_COMB.seq.*
- 91: /cgnl_6/ptodata/1/pna/US6016C_COMB.seq.*
- 92: /cgnl_6/ptodata/1/pna/US6017A_COMB.seq.*
- 93: /cgnl_6/ptodata/1/pna/US6017B_COMB.seq.*
- 94: /cgnl_6/ptodata/1/pna/US6017C_COMB.seq.*
- 95: /cgnl_6/ptodata/1/pna/US6018A_COMB.seq.*
- 96: /cgnl_6/ptodata/1/pna/US6018B_COMB.seq.*
- 97: /cgnl_6/ptodata/1/pna/US6018C_COMB.seq.*
- 98: /cgnl_6/ptodata/1/pna/US6019_COMB.seq.*
- 99: /cgnl_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
- 100: /cgnl_6/ptodata/1/pna/US06_NEW_COMB.seq.*
- 101: /cgnl_6/ptodata/1/pna/US07_NEW_COMB.seq.*
- 102: /cgnl_6/ptodata/1/pna/US08_NEW_COMB.seq.*
- 103: /cgnl_6/ptodata/1/pna/US09_NEW_COMB.seq.*
- 104: /cgnl_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69	100.0	69	28	US-08-935-377-1
2	57	82.6	57	28	US-08-935-377-5

Sequence 1, Appl1
Sequence 5, Appl1

```

3 53 76.8 53 28 US-08-935-377-23
4 53 76.8 53 28 US-08-935-377-24
5 49 71.0 145 28 US-08-935-377-6
6 49 71.0 148 28 US-08-935-377-7
7 48.8 70.7 149 28 US-08-935-377-8
8 48 69.6 150 28 US-08-935-377-9
9 37 53.6 75 28 US-08-935-377-3
10 36 52.2 70 28 US-08-935-377-30
11 32 46.4 71 28 US-08-935-377-31
12 28.6 41.4 40 12 US-08-301-013-27
13 28.6 41.4 40 17 US-08-592-610-406
14 28.6 41.4 40 47 US-09-295-309A-406
15 28.6 41.4 234 13 US-08-358-928-89
16 28.6 41.4 234 18 US-08-651-472-89
17 28.6 41.4 2360 12 US-08-301-013-30
18 28.6 41.4 2360 17 US-08-592-610-409
19 28.6 41.4 2360 47 US-09-295-309A-409
20 28.6 41.4 4987 12 US-08-301-013-31
21 28.6 41.4 4987 17 US-08-592-610-410
22 28.6 41.4 4987 47 US-09-295-309A-410
23 27.6 40.0 40 47 US-09-295-309A-411
24 27.2 39.4 1341 57 US-09-514-000-7200
25 27.2 39.4 4767 57 US-09-514-000-380
26 26.6 38.6 14103 86 US-06-144-351-2441
27 26.4 38.3 40 3 US-07-722-860-6
28 26.4 38.3 40 9 US-08-213-449A-6
29 26.4 38.3 40 9 US-08-213-449B-6
30 26.4 38.3 101 37 US-09-055-293-3
31 26.2 38.0 678 40 US-09-107-532-2649
32 26.2 38.0 7206 69 US-06-050-444-687
33 26.2 38.0 7206 70 US-06-068-186-691
34 26 37.7 870 16 US-08-474-135-3
35 26 37.7 870 19 US-08-708-856-3
36 26 37.7 870 55 US-09-455-406-3
37 26 37.7 1080 16 US-08-708-856-13
38 26 37.7 1081 16 US-08-474-135-11
39 26 37.7 1081 55 US-09-455-406-13
40 25.4 36.8 40 3 US-07-722-860-5
41 25.4 36.8 40 9 US-08-213-449A-5
42 25.4 36.8 40 9 US-08-213-449B-5
43 25.4 36.8 95 9 US-08-213-449B-20
44 25 36.2 40 16 US-08-479-869-8
45 25 36.2 40 16 US-08-484-575-8

ALIGNMENTS

RESULT 1
US-08-935-377-1
; Sequence 1, Application US/08935377
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; TITLE OF INVENTION: T Cells Specific for Target Antigens and
; TITLE OF INVENTION: Vaccines Based Thereon
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/935,377
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-935-377-5

Query Match 100.0%; Score 69; DB 28; Length 69;
Best Local Similarity 100.0%; Pred. No. 5.2e-15;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCAAAATTGAAACTAGATCTATTTATTCACGCGCGCCGCGCATGGCGCGCGGCC 60
Db 1 GGCCAAAATTGAAACTAGATCTATTTATTCACGCGCGCCGCGCATGGCGCGCGGCC 60
QY 61 AACGGCGGA 69
Db 61 AACGGCGGA 69

US-08-935-377-1
; NAME: Steffe, Eric K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 46..69
; US-08-935-377-1
```

```
Query Match      82.6%; Score 57; DB 28; Length 57;
Best Local Similarity 100.0%; Pred. No. 9.8e-11;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAAAATTCAGAACTAGATCTATTATTGTCACGCGCGCCGCGCCATGGCGCCGCC 57
    |||
DB 1 GCCCAAAATTCAGAACTAGATCTATTATTGTCACGCGCGCGCCGCGCCATGGCGCCGCC 57

RESULT 3
US-08-935-377-23
; Sequence 23, Application US/08935377
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; TITLE OF INVENTION: T Cells Specific for Target Antigens and
; TITLE OF INVENTION: Vaccines Based Thereon
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,377
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-935-377-23

Query Match      76.8%; Score 53; DB 28; Length 53;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAAAATTCAGAACTAGATCTATTATTGTCACGCGCGCCGCGCCATGGCGCCGCC 53
    |||
DB 1 GCCCAAAATTCAGAACTAGATCTATTATTGTCACGCGCGCGCCGCGCCATGGCGCCGCC 53

RESULT 4
US-08-935-377-24/c
; Sequence 24, Application US/08935377
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; TITLE OF INVENTION: T Cells Specific for Target Antigens and
; TITLE OF INVENTION: Vaccines Based Thereon
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,377
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-935-377-23

Query Match      76.8%; Score 53; DB 28; Length 53;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAAAATTCAGAACTAGATCTATTATTGTCACGCGCGCGCCGCGCCATGGCGCCGCC 53
    |||
DB 1 GCCCAAAATTCAGAACTAGATCTATTATTGTCACGCGCGCGCCGCGCCATGGCGCCGCC 53

RESULT 5
US-08-935-377-6
; Sequence 6, Application US/08935377
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; TITLE OF INVENTION: T Cells Specific for Target Antigens and
; TITLE OF INVENTION: Vaccines Based Thereon
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,377
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-935-377-24

Query Match      76.8%; Score 53; DB 28; Length 53;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AAAAATTGAAAACTAGATCTATTATTGTCACGCGCGCCGCGCCATGGCGCCGCC 57
    |||
DB 53 AAAAATTGAAAACTAGATCTATTATTGTCACGCGCGCGCCGCGCCATGGCGCCGCC 1

RESULT 5
US-08-935-377-6
; Sequence 6, Application US/08935377
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; TITLE OF INVENTION: T Cells Specific for Target Antigens and
; TITLE OF INVENTION: Vaccines Based Thereon
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,377
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```


TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-935-377-9

Query Match 69.6%; Score 48; DB 28; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAAAATGAAAACTAGATCTATTATTGCACGGCGCCGCATG 48
|||||
Db 1 GCCCAAAATGAAAACTAGATCTATTATTGCACGGCGCCGCATG 48

RESULT 9

US-08-935-377-3
Sequence 3, Application US/08935377
GENERAL INFORMATION:
APPLICANT: Zauderer, Maurice
TITLE OF INVENTION: T Cells Specific for Target Antigens and
TITLE OF INVENTION: Vaccines Based Thereon
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/935,377
FILING DATE: 22-SEP-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 75 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 52..75
US-08-935-377-3

Query Match 53.6%; Score 37; DB 28; Length 75;
Best Local Similarity 78.7%; Pred. No. 0.0016;
Matches 59; Conservative 0; Mismatches 10; Indels 6; Gaps 1;

QY 1 GCCCAAAATGAAAACTAGATCTATTATTG-----CACGGCGCGCCCATGGGCCG 54
|||||
Db 1 GCCCAAAATGAAATTTATTTTITTTTGGTAATAAGCGCGCCCATGGGCCG 60

QY 55 GCCGCCAACGGCGGA 59
|||||
Db 61 GCCGCCAACGGCGGA 75

RESULT 10

US-08-935-377-30
Sequence 30, Application US/08935377
GENERAL INFORMATION:
APPLICANT: Zauderer, Maurice
TITLE OF INVENTION: T Cells Specific for Target Antigens and
TITLE OF INVENTION: Vaccines Based Thereon
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/935,377
FILING DATE: 22-SEP-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 70 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-935-377-30

Query Match 52.2%; Score 36; DB 28; Length 70;
Best Local Similarity 100.0%; Pred. No. 0.0035;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAAAATGAAAACTAGATCTATTATTATGCAC 36
|||||
Db 1 GCCCAAAATGAAAACTAGATCTATTATTATGCAC 36

RESULT 11

US-08-935-377-31/c
Sequence 31, Application US/08935377
GENERAL INFORMATION:
APPLICANT: Zauderer, Maurice
TITLE OF INVENTION: T Cells Specific for Target Antigens and
TITLE OF INVENTION: Vaccines Based Thereon
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/935,377
FILING DATE: 22-SEP-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 71 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-935-377-31

Query Match 46.4%; Score 32; DB 28; Length 71;
Best Local Similarity 100.0%; Pred. No. 0.095;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AAAAAATTGAAAACTAGATCTATTTATTCAC 36
|||||
DB 71 AAAAAATTGAAAACTAGATCTATTTATTCAC 40

RESULT 12
US-08-301-013-27
; Sequence 27, Application US/08301013
; GENERAL INFORMATION:
; APPLICANT: Makoto SEKI et al.
; TITLE OF INVENTION: ANTIVIRAL AGENT
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/301,013
; FILING DATE: September 6, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/299,113
; FILING DATE: September 2, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid synthetic DNA

US-08-301-013-27

Query Match 41.4%; Score 28.6; DB 12; Length 40;
Best Local Similarity 88.6%; Pred. No. 1.3;
Matches 31; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CCAAAATTGAAAACTAGATCTATTTATTCACG 37
|||||
DB 4 CCAAAATTGAAAACTAGATCTATTTATTCACG 38

RESULT 13
US-08-592-610-406
; Sequence 406, Application US/08592610
; GENERAL INFORMATION:
; APPLICANT: Makoto SEKI et al.
; TITLE OF INVENTION: ANTISENSE COMPOUNDS COMPLEMENTARY TO HCV
; TITLE OF INVENTION: GENOME
; NUMBER OF SEQUENCES: 410
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,610
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/111,465
; FILING DATE: August 25, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew Jacob
; REGISTRATION NUMBER: 25,154
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; INFORMATION FOR SEQ ID NO: 406:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other nucleic acid, synthetic DNA
US-08-592-610-406

Query Match 41.4%; Score 28.6; DB 17; Length 40;
Best Local Similarity 88.6%; Pred. No. 1.3;
Matches 31; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CCAAAATTGAAAACTAGATCTATTTATTCACG 37
|||||
DB 4 CCAAAATTGAAAACTAGATCTATTTATTCACG 38

RESULT 14
US-09-295-309A-406
; Sequence 406, Application US/09295309A
; GENERAL INFORMATION:
; APPLICANT: Makoto SEKI et al.
; TITLE OF INVENTION: ANTISENSE COMPOUNDS COMPLEMENTARY TO HCV
; TITLE OF INVENTION: GENOME
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
STREET: 2033 K Street, N.W., Suite 800
CITY: Washington
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/295,309A
FILING DATE: April 21, 1999
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Matthew Jacob
REGISTRATION NUMBER: 25,154
REFERENCE/DOCKET NUMBER:
TELEPHONE: (202) 721-8200
TELEFAX: (202) 721-8250
INFORMATION FOR SEQ ID NO: 406:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid, synthetic DNA
US-09-295-309A-406

Query Match 41.4%; Score 28.6; DB 47; Length 40;
Best Local Similarity 88.6%; Pred. No. 1.3;
Matches 31; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 CCAAAATTGAAAACTAGATCTATTATTGACG 37
|||||
DB 4 CCAAAATTGAAAACTAGTCTAATTATTGACG 38
|||||

RESULT 15

US-08-358-928-89
Sequence 89, Application US/08358928
GENERAL INFORMATION:
APPLICANT: DORNER, Friedrich
APPLICANT: SCHEIFLINGER, Friedrich
APPLICANT: FALKNER, Falko Gunter
APPLICANT: FLEIDERER, Michael
TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF CHIMERIC
TITLE OF INVENTION: VIRUSES CONTAINING HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/358,928
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/914,738
FILING DATE: 20-JUL-1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/750,080
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/166/IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 234 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-358-928-89

Query Match 41.4%; Score 28.6; DB 13; Length 234;
Best Local Similarity 72.5%; Pred. No. 2.2;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 5 AAAAAATTGAAAACTAGATCTATTATTGACGCGCGCCCATGGCGCGG 55
|||||
DB 183 AAAAAATTGAAAACTATTCTAATTATTGACGCGGTACGTACCATGGCGCGG 233
|||||

Search completed: May 30, 2000, 09:48:19
Job time: 60169 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 11:04:41 ; Search time 2192.43 Seconds
(without alignments)
127.563 Million cell updates/sec

Title: US-08-935-377-1
Perfect score: 69
Sequence: 1 GCCCAAAATGAAAACTA.....GCCCGGGCCGCAACGGCGGA 69

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: em_est20:*
50: em_est21:*
51: em_est22:*
52: gb_est30:*
53: gb_est31:*
54: gb_est32:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: em_gss5:*
93: em_gss6:*
94: gb_gss7:*
95: gb_gss8:*
96: gb_gss9:*
97: em_gss7:*
98: em_gss8:*
99: gb_gss11:*
100: gb_gss10:*
101: em_gss9:*
102: em_gss10:*
103: em_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

AA060654/c
LOCUS AA060654 464 bp mRNA EST 30-SEP-1997
DEFINITION vm93a06.r1 Knowles Solter mouse blastocyst B1 Mus musculus cDNA clone IMAGE:1005778 5' similar to gb:L33715 Mus musculus Fau gene, complete cds (MOUSE);, mRNA sequence.

ACCESSION AA060654
VERSION AA060654.1 GI:2455588
SOURCE EST.
ORGANISM house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 464)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
TITLE On Apr 14, 1993 this sequence version replaced gi:693296.
JOURNAL Contact: Marra M/Mouse EST Project
COMMENT WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:569994
 Putative full length read
 vector to vector length is 491.
FEATURES Location/Qualifiers
 source
 1..464
 /organism="Mus musculus"
 /strain="B6D2 F1/J"
 /db_xref="taxon:10090"
 /clone_lib="IMAGE:1005778"
 /tissue_type="blastocyst"
 /dev_stage="embryo (pre-implantation)"
 /lab_host="DH10B"
 /note="Organ: embryo; Vector: pSPORT; Site_1: NotI; Site_2: SalI; Cloned unidirectionally from mRNA prepared from 800 blastocysts. Primer: SalI(dT): 5'-CGGTGACCGTCGACCGGTTTCTTTT-3'. cDNAs were cloned into the NotI/SalI sites of a pSPORT vector (Life Technologies). Two different size selections: BI (larger inserts) and B3."
 BASE COUNT 114 a 131 c 141 g 78 t
 ORIGIN

Query Match 38.6%; Score 26.6; DB 36; Length 464;
 Best Local Similarity 66.7%; Pred. No. 76; Indels 0; Gaps 0;
 Matches 38; Conservative 0; Mismatches 19;

QY 1 GGCCAAAATTTGAAAACATAGATCTATTATTGACGCGCGCCATGGCGCCGCC 57
 ||| || ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 407 GGCACACATGACAAGCGCGGTTTCTTACTGCATTCGGCCCTTCTGGCCGCC 351

RESULT 3
LOCUS AQ274419 692 bp DNA GSS 03-NOV-1998
DEFINITION mgxb0015M12r CUGI Rice Blast BAC Library Pyricularia grisea genomic clone mgxb0015M12r, genomic survey sequence.
ACCESSION AQ274419
VERSION AQ274419.1 GI:3827734
KEYWORDS GSS.
SOURCE Pyricularia grisea.
ORGANISM Pyricularia grisea
 Eukaryota; Fungi; Ascomycota; anamorphic Ascomycota; Pyricularia.
 1 (bases 1 to 692)
 Yu,Y., Zhu,H., Boyd,C.A., Gaudette,B., Gayle,A., Kingsbury,R., Phillips,K., Sasnowski,M., Wing,R.A. and Dean,R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome
 Unpublished (1998)
TITLE Contact: Dean RA
JOURNAL Clemson University Genomics Institute
COMMENT Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: GGAACACAGTATGACCATG
 Class: BAC ends
 High quality sequence stop: 390.
FEATURES Location/Qualifiers
 source
 1..692
 /organism="Pyricularia grisea"
 /strain="70-15"
 /db_xref="taxon:89476"
 /clone_lib="mgxb0015M12r"
 /clone_lib="CUGI Rice Blast BAC Library"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."
 BASE COUNT 145 a 169 c 210 g 167 t
 ORIGIN

Query Match 37.4%; Score 25.8; DB 105; Length 692;
 Best Local Similarity 67.9%; Pred. No. 1.3e+02;
 Matches 36; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 8 AATTGAAAACTAGATCTATTATTGACGCGCGCCATGGCGCCGCCGCC 60
 ||| || ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 59 AATGGACCGCTAGGCTCTTCTTGTGACGCGCGCCCTTGGCCTTGCACCC 111

RESULT 4
LOCUS CNS0164L 1201 bp DNA GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC BACN15C21 of DrosBAC library from Drosophila melanogaster (fruit fly); genomic survey sequence.
ACCESSION AL106287
VERSION AL106287.1 GI:5621177
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1201)
 Direct Submission
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
JOURNAL - Web : www.genoscope.cns.fr
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC

library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.

FEATURES

source

Location/Qualifiers
1..1201
/organism="Drosophila melanogaster"
/plasmid="pBeloBAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN15C21"
/note="end : Sp6"

BASE COUNT 304 a 204 c 230 g 275 t 188 others
ORIGIN

Query Match 37.1%; Score 25.6; DB 83; Length 1201;
Best Local Similarity 32.8%; Pred. No. 1.4e+02;
Matches 20; Conservative 26; Mismatches 15; Indels 0; Gaps 0;

QY 4 CAATAAATGAAACATGATCTATTATTGACGCGCGCCATGGCGCGCCGCCAAC 63
Db 847 CAATAAAMAAAAHHWWHHWWTTTTCSSSGSSSVGKGACCCSCCNSCS 906

QY 64 G 64
Db 907 G 907

RESULT 5

LOCUS

DEFINITION HS_5103_B2_G10_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=679 Col=20 Row=N, genomic survey sequence.

ACCESSION AQ406794

VERSION 1

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL PROC. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University Of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallaceu.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC resources (http://bacpac.med.buffalo.edu/ordering.bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 679 row: N column: 20
Seq primer: T7
Class: BAC ends
High quality sequence stop: 492.

FEATURES

source

Location/Qualifiers
1..492
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="plate=679 Col=20 Row=N"
/clone_lib="RPCI-11 Human Male BAC Library"

Query Match 36.2%; Score 25; DB 99; Length 507;
Best Local Similarity 69.4%; Pred. No. 2.4e+02;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 2 GCCAAAATGAAACATGATCTATTATTGACGCGCGCCGCCATGGG 50
Db 154 GCCAATATCTGAAGATATCTATTATATGCTATCGTCGCCATAGG 202

BASE COUNT 142 a 116 c 112 g 137 t
ORIGIN

FEATURES

source

Location/Qualifiers
1..507
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=3115 Col=22 Row=P"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 142 a 116 c 112 g 137 t

/sex="male"
/note="vector: pBACe3.6; Genomic sequence of BAC ends"
BASE COUNT 170 a 96 c 93 g 131 t 2 Others
ORIGIN

Query Match 36.2%; Score 25; DB 106; Length 492;
Best Local Similarity 61.5%; Pred. No. 2.4e+02;
Matches 40; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 GGCACAAAATGAAACATGATCTATTATTGACGCGCGCCGCCATGGCGCGCGCC 60
Db 239 GGACATATCTGGATATTGAGAACTCAATGCTGAACAAGCTGACATGCGCCCTGCCTC 298

QY 61 AACGG 65

Db 299 AATGG 303

RESULT 6

LOCUS

DEFINITION AQ212765

507 bp DNA GSS

HS_3115_B2_H11_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3115 Col=22 Row=P, genomic survey sequence.

ACCESSION AQ212765

VERSION 1

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

COMMENT On Sep 10, 1998 this sequence version replaced gi:3553850.
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallaceu.washington.edu
Sequence Tagged Connector
Plate: 3115 row: P column: 22
Class: BAC ends
High quality sequence stop: 507.

```

RESULT 7
LOCUS   AQ699107          516 bp    DNA          GSS          06-JUL-1999
DEFINITION   HS_5560_B2_E01_77A RPCI-11 Human Male BAC Library Homo sapiens
              genomic clone Plate-1136 Col=2 Row=J, genomic survey sequence.
ACCESSION   AQ699107
VERSION     AQ699107.1  GI:5389355
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 516)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE       Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL     Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE     99380589
COMMENT     On Sep 10, 1998 this sequence version replaced gi:3555646.
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1136 row: J column: 2
            Seq primer: 17
            Class: BAC ends
FEATURES    High quality sequence stop: 516.
            Location/Qualifiers
                1..516
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="Plate-1136 Col=2 Row=J"
                /clone_lib="RPCI-11 Human Male BAC Library"
                /sex="male"
                /note="Vector: pBACE3.6; Genomic sequence of BAC ends"
BASE COUNT 173 a 103 c 107 g 127 t 6 Others
ORIGIN
Query Match 36.2%; Score 25; DB 85; Length 516;
Best Local Similarity 61.5%; Pred. No. 2.4e+02;
Matches 40; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1 GCCCAAAATTGAAAACATAGATCTATTATTTCACGCGCGCCGATGGCGCCGCGCCG 60
    || || || || || || || || || || || || || || || || || || || ||
Db 236 GGACATATACGTGATTTGAGAACTCAATGGTGAACAAGGCTGACATGCGCCCTGCCCTC 295
QY 61 AACGG 65
    || ||
Db 296 AATGG 300

RESULT 8
LOCUS   AQ415486          603 bp    DNA          GSS          23-MAR-1999
DEFINITION   RPCI-11-177019.TJ RPCI-11 Homo sapiens genomic clone
              RPCI-11-177019, genomic survey sequence.
ACCESSION   AQ415486
VERSION     AQ415486.1  GI:4474455
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 603)
AUTHORS     Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
TITLE       Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
JOURNAL     Unpublished (1997)
COMMENT     Other GSSs: RPCI11-122P9.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic genomics

```

```

ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 603)
AUTHORS     Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
TITLE       Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
JOURNAL     Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbs@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES    Location/Qualifiers
                1..603
                /organism="Homo sapiens"
                /db_xref="GDB:7567938"
                /db_xref="taxon:9606"
                /clone="RPCI-11-177019"
                /clone_lib="RPCI-11"
                /sex="Male"
                /cell_type="Lymphocytes"
                /note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
                RPC111 Human Male BAC Library"
BASE COUNT 219 a 116 c 120 g 148 t
ORIGIN
Query Match 36.2%; Score 25; DB 106; Length 603;
Best Local Similarity 61.5%; Pred. No. 2.4e+02;
Matches 40; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1 GCCCAAAATTGAAAACATAGATCTATTATTTCACGCGCGCCGATGGCGCCGCGCCG 60
    || || || || || || || || || || || || || || || || || || || ||
Db 237 GGACATATACGTGATTTGAGAACTCAATGGTGAACAAGGCTGACATGCGCCCTGCCCTC 296
QY 61 AACGG 65
    || ||
Db 297 AATGG 301

RESULT 9
LOCUS   AQ384188          658 bp    DNA          GSS          21-MAY-1999
DEFINITION   RPCI11-122P9.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-122P9,
              genomic survey sequence.
ACCESSION   AQ384188
VERSION     AQ384188.1  GI:4355211
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 658)
AUTHORS     Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
TITLE       Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
JOURNAL     Unpublished (1997)
COMMENT     Other GSSs: RPCI11-122P9.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic genomics

```

The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES

Location/Qualifiers
1..658
/organism="Homo sapiens"
/db_xref="GDB:7546832"
/db_xref="taxon:9606"
/clone="RPCI-11-122P9"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"

BASE COUNT 235 a 128 c 126 g 169 t
ORIGIN

Query Match 36.2%; Score 25; DB 106; Length 658;
Best Local Similarity 61.5%; Pred. No. 2.4e+02;
Matches 40; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 1 GCCCAAAATTGAAACTAGATCTATTATTGCACGGCGCCGATGGCGCCGCCGCC 60
Db 225 GCATATACCTGATTTGAGAACTCAATGTTGAACAGGTGATCGCCCTGCCCTC 284

Qy 61 AACG 65
Db 285 AATGG 289

RESULT 10
AW265066/c
LOCUS
DEFINITION
xq60g03.x1 NCI_CGAP_Co22 Homo sapiens cDNA clone IMAGE:2755060 3'
similar to SW:E411_ADE02 P03241 PROBABLE EARLY E4 11 KD PROTEIN.
[1] ; mRNA sequence.

ACCESSION AW265066.1 GI:6641882
VERSION AW265066
KEYWORDS EST.
SOURCE human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 426)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT On Jul 7, 1999 this sequence version replaced gi:5406301.
Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
unknown library type

Possible reversed clone: similarity on wrong strand
Seq primer: -40UP from Gibco

High quality sequence stop: 425.
Location/Qualifiers

FEATURES

Location/Qualifiers
1..426
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2755060"

/clone_lib="NCI_CGAP_Co22"
/tissue_type="colonic adenocarcinoma"
/lab_host="DH10B"

/note="Organ: colon; Vector: pAMP10; cDNA made by oligo-dT priming. Non-directionally cloned into the UDG sites of PAMP10. Size-selected on agarose gel, average insert size 500 bp. Primary library; non-amplified. cDNA Library Preparation: David B. Krizman, Ph.D (NCI). Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 103 a 80 c 118 g 124 t
ORIGIN

Query Match 35.9%; Score 24.8; DB 79; Length 426;
Best Local Similarity 63.3%; Pred. No. 2.8e+02;
Matches 38; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 3 CCAAAATGAAACTAGATCTATTATTGCACGGCGCCGATGGCGCCGCCGCCAA 62
Db 257 CCAAACTCAAAATGAAGATCTATTAACTGAACGGCTCCCTCCGGTGGTGTCAA 198

RESULT 11
AW276836/c

LOCUS
DEFINITION
xp66g11.x1 NCI_CGAP_Ov39 Homo sapiens cDNA clone IMAGE:2745380 3'
similar to SW:E413_ADE02 P03240 PROBABLE EARLY E4 13 KD PROTEIN. ;
mRNA sequence.

ACCESSION AW276836
VERSION AW276836.1 GI:6663866
KEYWORDS EST.
SOURCE human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 630)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT On Nov 22, 1999 this sequence version replaced gi:6462296.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert_Strausberg@nih.gov
Tissue Procurement: Andrew Berchuck M.D., John Gillespie M.D.,
Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: similarity on wrong strand
Possible reversed clone: polyt not found

Seq primer: -40UP from Gibco
High quality sequence stop: 422.

FEATURES

Location/Qualifiers
1..630
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2745380"
/clone_lib="NCI_CGAP_Ov39"
/sex="female"

/tissue_type="papillary serous ovarian metastasis"
/lab_host="DH10B"

/note="Organ: ovary; Vector: pAMP10; cDNA made by oligo-dT priming. Non-directionally cloned into the UDG sites of PAMP10. Size-selected on agarose gel, average insert size 500 bp. Primary library; non-amplified. cDNA Library Preparation: David B. Krizman, Ph.D (NCI). Reference: Krizman et al. (1996) Cancer Research


```

SOURCE          Pig.
ORGANISM        Sus scrofa
REFERENCE
AUTHORS         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
TITLE          1 (bases 1 to 194)
JOURNAL        Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
                Stone, R.I., Heaton, M.P., Grosse, W.M., Bennett, G.A., and Keeler, J.W.
                Design and use of two pooled tissue normalized cDNA libraries for
                EST discovery in swine
COMMENT        Unpublished (2000)
                On May 18, 1998 this sequence version replaced gi:3138121.
                Contact: Smith TPL
                USDA, ARS, US Meat Animal Research Center
                PO Box 166, Clay Center, NE 68933-0166, USA
                Tel: 402 762 4366
                Fax: 402 762 4390
                Email: smitht@mail.marc.usda.gov
                Single pass sequencing. Bases called and trimmed with phred
                v0.980904.e. Vector identified by cross_match with the -minscore 20
                and -minmatch 12 options.
                PCR Primers
                FORWARD: AGGAACAGCTATGACCAT
                BACKWARD: GTTTCCTCAGTCACGAG
                Plate: 13 row: A column: 24
                Seq primer: ATTAGGTGACACTATAG.
                Location/Qualifiers
                source
                1..194
                /organism="Sus scrofa"
                /db_xref="taxon:9823"
                /clone_lib="MARC 2Pig"
                /tissue_type="pooled"
                /lab_host="DHI08"
                /note="vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                Library made from pooled tissue from testis, ovary,
                endometrium, hypothalamus, pituitary, and placenta."
                33 a 50 c 93 g 18 t

BASE COUNT      33 a 50 c 93 g 18 t
ORIGIN
Query Match      35.4%; Score 24.4; DB 80; Length 194;
Best Local Similarity 63.8%; Pred. No. 4.1e+02;
Matches 37; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 11 TGAAGAGCGAGGTGGAGGCTCGGAGCGCGCGCCATGGCCGCGCCGCCAAGCGCGG 68
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21 TGAAGAGCGAGGTGGAGGCTCGGAGCGCGCGCCATGGCCGCGCGCCGCCAAGCGCGG 78
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
AQ064961/c      355 bp DNA GSS 04-AUG-1998
LOCUS           HS_2225_Al_G06_MF CIT Approved Human Genomic Sperm Library D Homo
DEFINITION      sapiens genomic clone Plate=2225 Col=11 Row=M, genomic survey
                sequence
ACCESSION       AQ064961
VERSION         AQ064961.1 GI:3380789
KEYWORDS        GSS.
SOURCE          human.
ORGANISM        Homo sapiens
REFERENCE       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS         Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
                Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and
                Hood, L.
TITLE           Sequence-tagged connectors: A sequence approach to mapping and
                scanning the human genome
JOURNAL         Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE        99380589.
COMMENT         Contact: Mahairas GG, Wallace JC, Hood L
                High Throughput Sequencing Center
                University of Washington
                401 Queen Anne Avenue North, Seattle, WA 98109, USA

```

```

BASE COUNT      154 a 134 c 173 g 168 t 1 others
ORIGIN
Query Match      35.9%; Score 24.8; DB 79; Length 630;
Best Local Similarity 63.3%; Pred. No. 2.7e+02;
Matches 38; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 3 CCAAAAATTGAAAACCTAGATCTATTATTGACGCGCGCGCCATGGCCGCGCGCCCAA 62
    ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Db 290 CCARACCTCAAGATGATCTATTAGTGAACGCGCTCCCTCCGTTGGTGTCACAA 231
    ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||

RESULT 12
C84505          379 bp mRNA EST 26-MAR-1999
LOCUS           C84505 osteoclast subtracted library Oryctolagus cuniculus cDNA,
DEFINITION      mRNA sequence.
ACCESSION       C84505
VERSION         C84505.1 GI:4527765
KEYWORDS        EST.
SOURCE          Oryctolagus cuniculus.
ORGANISM        Oryctolagus cuniculus.
REFERENCE       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Lagomorpha; Leporidae; Oryctolagus.
AUTHORS         1 (bases 1 to 379)
                Kobori, M., Ikeda, Y., Nara, H., Kato, M., Kumegawa, M., Nojima, H. and
                Kawashima, H.
TITLE           Large scale isolation of osteoclast-specific genes by an improved
                method involving the preparation of a subtracted cDNA library
JOURNAL         Genes Cells 3 (7), 459-475 (1998)
MEDLINE        98424349
COMMENT         On Jun 5, 1998 this sequence version replaced gi:3189094.
                Contact: Kobori M
                Molecular Medicine Laboratories
                Institute for Drug Discovery Research, Yamanouchi Pharmaceutica
                21, Miyukigaoka, Tsukuba, Ibaraki 305, Japan
                Email: kobori@yamanouchi.co.jp
                PROJECT = "OSG".
                Location/Qualifiers
                source
                1..379
                /organism="Oryctolagus cuniculus"
                /db_xref="taxon:9986"
                /clone_lib="osteoclast subtracted library"
                /tissue_type="long bone"
                /cell_type="osteoclast"
                /cell_line="primary"
                /dev_stage="5 day-old"
                72 a 125 c 108 g 58 t 16 others

BASE COUNT      72 a 125 c 108 g 58 t 16 others
ORIGIN
Query Match      35.7%; Score 24.6; DB 48; Length 379;
Best Local Similarity 60.9%; Pred. No. 3.3e+02;
Matches 39; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 1 GGCCAAAATTGAAAACCTAGATCTATTATTGACGCGCGCGCCATGGCCGCGCGCCGCC 60
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 33 GGACCAACCTCAGAACCTGGACCCATTTGNTGACCTGGCGGACCTCAGCTCCGCGCTCC 92
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 61 AACG 64
    || |
Db 93 AAGG 96

RESULT 13
AW354184        194 bp mRNA EST 02-FEB-2000
LOCUS           AW354184
DEFINITION      32589 MARC 2Pig Sus scrofa cDNA 5', mRNA sequence.
ACCESSION       AW354184
VERSION         AW354184.1 GI:6853174
KEYWORDS        EST.

```

Wed May 31 10:04:51 2000

Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2225 row: M column: 11
Class: BAC ends
High quality sequence stop: 355.
Location/Qualifiers

FEATURES
source

1. .355
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="plate=2225 Col=11 Row=M"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 111 a 70 c 71 g 103 t

ORIGIN

Query Match 35.4%; Score 24.4; DB 96; Length 355;
Best Local Similarity 73.8%; Pred. No. 3.9e+02;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 3 CCAAAATGAAACTAGATCTATTATTCACGCGCGCCG 44

DB 197 CAAGAATTAAACAAGTAGATGTATCTATTGCAAGCAGCC 156

RESULT 15

C78547/c

LOCUS

DEFINITION

C78547 Mouse 3.5-dpc blastocyst cDNA Mus musculus cDNA clone

J0051B03 3' similar to Mus musculus transcription factor like

protein 4 TCFL4, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 578)
Ko, M.S.H., Kitchen, J.R., Wang, X., Wang, X., Threat, T.A., Sun, T.,
Grahovac, M.J., Mason, S., Lim, M.K., Paonessa, P.D., Sauls, A.D. and
Doi, H.

Systematic analyses of genes expressed in 3.5-dpc mouse blastocyst

(The ERATO/Doi Project at Wayne State University)

Unpublished (1997)

On May 5, 1995 this sequence version replaced gi:797882.

JOURNAL

COMMENT

Contact: Hirofumi Doi
Doi Biosymmetry Project, ERATO
Japan Science and Technology Corporation (JST)
WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan
Email: hdbioa.jst.go.jp.

FEATURES

source

1. .578
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="J0051B03"
/clone_lib="Mouse 3.5-dpc blastocyst cDNA"
/tissue_type="blastocyst"
/dev_stage="3.5-dpc"
Location/Qualifiers
128 a 128 c 152 g 128 t 2 others

BASE COUNT

ORIGIN

Query Match 35.4%; Score 24.4; DB 36; Length 578;
Best Local Similarity 62.7%; Pred. No. 3.7e+02;
Matches 37; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 4 CAAAAATGAAACTAGATCTATTATTCACGCGCGCCATGGCGCCGCCGCAA 62

Db 448 CATCAGTTGAAAAACCAACTCTACTATGATTCGCCATGGAAGCCNGCAGAGCCCAACAA 390

Search completed: May 29, 2000, 21:13:10
Job time: 36509 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:33:53 ; Search time 1214.87 Seconds
(without alignments)
470.244 Million cell updates/sec

Title: US-08-935-377-3

Perfect score: 75
Sequence: 1 GCCCAAAATGAAATTTA.....GCCCGCGCGCAACGGCGGA 75

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, 3808571567 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pli:*
8: gb_pli2:*
9: gb_pri:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vi:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pi:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vi:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*

44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	40.6	54.1	2164	16	VVH3FDP1	X89856 V.virus DNA
2	40.6	54.1	2164	16	VVH3FDP2	X89857 V.virus DNA
3	40	53.3	84	5	A59039	A59039 Sequence 27
4	40	53.3	86	5	A59040	A59040 Sequence 28
5	32.4	43.2	160821	54	AC012468	AC012468 Homo sapi
6	31.8	42.4	331668	41	AC009246	AC009246 Homo sapi
7	31	41.3	4376	12	MUSIL13A	L13028 Mouse inter
8	31	41.3	159500	12	AC005742	AC005742 Mus muscu
9	31	41.3	196382	55	AC020886	AC020886 Mus muscu
10	30.4	40.5	201746	41	AC006895	AC006895 Caenorhab
11	30.2	40.3	161449	40	AC005341	AC005341 Homo sapi
12	30.2	40.3	161866	45	AC021672	AC021672 Homo sapi
13	29.8	39.7	108061	8	AC009526	AC009526 Arabidops
14	29.8	39.7	141655	32	HL1093G12	HL121751 Homo sapi
15	29.8	39.7	218305	33	HSJ324O17	AL110115 Homo sapi
16	29.8	39.7	218305	33	HSJ324O17	AL110115 Homo sapi
17	29.6	39.5	62944	40	HSMHC3W5A	U89336 Homo sapien
18	29.6	39.5	183285	42	AC009833	AC009833 Homo sapi
19	29.6	39.5	275670	55	AC022328	AC022328 Mus muscu
20	29.4	39.2	35184	42	AC014624	AC014624 Drosophil
21	29.4	39.2	39573	35	CELC08D9	AF101319 Caenorhab
22	29.2	38.9	35974	34	CELC46F1	U53148 Caenorhabdi
23	29.2	38.9	134095	41	AC006915	AC006915 Caenorhab
24	29.2	38.9	142232	56	AC008443	AC008443 Homo sapi
25	29.2	38.9	179100	42	AC012641	AC012641 Homo sapi
26	29.2	38.9	227066	41	AC006877	AC006877 Caenorhab
27	29.2	38.9	298804	41	AC006911	AC006911 Caenorhab
28	29	38.7	31531	34	CEZC434	Z75714 Caenorhabdi
29	29	38.7	73461	56	AC010563	AC010563 Drosophil
30	29	38.7	110000	32	CXY106G6_5	Continuation (6 of
31	29	38.7	164197	44	AC016328	AC016328 Homo sapi
32	29	38.7	191258	45	AC008462	AC008462 Homo sapi
33	28.8	38.4	61450	10	HS15D7	AL031229 Human DNA
34	28.8	38.4	96308	40	AF190465	AF190465 Homo sapi
35	28.8	38.4	149259	43	AC013668	AC013668 Homo sapi
36	28.8	38.4	151595	54	AC015809	AC015809 Homo sapi
37	28.8	38.4	165093	41	AC005517	AC005517 Homo sapi
38	28.8	38.4	256172	41	AC005139	AC005139 Plasmodiu
39	28.6	38.1	2166	35	AF200327	AF200327 Plasmodiu
40	28.6	38.1	27885	34	CELC488	AF022984 Caenorhab
41	28.6	38.1	30452	9	AB014077	AB014077 Homo sapi
42	28.6	38.1	38561	41	AC006604	AC006604 Caenorhab
43	28.6	38.1	43600	11	AC004209	AC004209 Homo sapi
44	28.6	38.1	100000	9	AP000514	AP000514 Homo sapi
45	28.6	38.1	101584	41	AC006793	AC006793 Caenorhab

ALIGNMENTS

Wed May 31 10:04:51 2000

```

RESULT 1
VH3FDPCL1      2164 bp      DNA      VRL      03-JAN-1996
LOCUS          V.virus DNA for Hind III F fragment (clone VDPCL1).
DEFINITION
ACCESSION      X89856
VERSION        X89856.1 GI:927568
KEYWORDS       HindIII F fragment; hph gene; hygromycin phosphotransferase;
SOURCE         Vaccinia virus.
ORGANISM       Vaccinia virus
VIRUSES; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
Orthopoxvirus.
1 (bases 1 to 2164)
Pfleiderer, M., Falkner, F.G. and Dörner, F.
A novel vaccinia virus expression system allowing construction of
recombinants without the need for selection markers, plasmids and
bacterial hosts
J. Gen. Virol. 76 (Pt 12), 2957-2962 (1995)
JOURNAL
MEDLINE        96112180
AUTHORS        Pfleiderer, M.
TITLE          Direct Submission
SUBMITTED      (20-JUL-1995) M. Pfleiderer, Immuno AG, A. 2304
ORTH/DONAU, Uferstrasse 15, AUSTRIA
COMMENT        Related sequence M35027.
FEATURES       Location/Qualifiers
1..2164
/organism="Vaccinia virus"
/strain="WR"
/species="taxon:"mammalia"
/db_xref="taxon:10245"
/lab_host="mammalian cell lines"
/cell_line="recombinant Vaccinia virus"
/clone="VDPCL1"
/map="45478 nt number"
1..485
/note="Hind III F fragment"
479..1662
/note="NOT I gene cassette"
479
/note="NOT I cleavage site corresponding to the unique NOT
I site in the Vv genome at position 45478 (strain
copenhagen)"
486..543
/note="strong synthetic early/late promoter"
536
/note="unique Sfi I cleavage site"
544..594
/note="weak synthetic early/late promoter"
595..1620
/gene="hph"
595..1620
/gene="hph"
/codon_start=1
/product="hygromycin phosphotransferase"
/protein_id="CAA61952.1"
/db_xref="GI:953210"
/db_xref="SWISS-PROT:P00557"
/translation="MKKPELTATSVKFLIEKFDVSDLMOLSEGESRAFSFDVGG
GYVLRVNSCADGFYKDRYVYRHFASALPIPEVLDIGEFSESLTYCISRRAGVTLOD
LPETEPALVQPAEAMDAADLSQTSQGFPGFQGTWTRDTCIALADPHY
HWQTVMDTVASVAQALDELMLWEDCEVRLHVFADFGSNVLTIDNGRTAVLDWS
EAMFGDSQYEVANIFFRWPLACMEQOTYFFRRHPFELAGSPRLRAYMLRIGLDOLYQ
SLVDGNFDDAAWAQGRCDALVRSGAGTVGRTQIARRSAVWTDGCVLEADSGNRRP
TPRAKE"
1635
/note="Rsr II cleavage site"
1642
/note="NOT I cleavage site"
1643..2164
/note="Hind III F fragment"

RESULT 2
VH3FDPCL2      2164 bp      DNA      VRL      03-JAN-1996
LOCUS          V.virus DNA for Hind III F fragment (clone VDPCL2).
DEFINITION
ACCESSION      X89857
VERSION        X89857.1 GI:927569
KEYWORDS       HindIII F fragment; hph gene; hygromycin phosphotransferase;
SOURCE         Vaccinia virus.
ORGANISM       Vaccinia virus
VIRUSES; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
Orthopoxvirus.
1 (bases 1 to 2164)
Pfleiderer, M., Falkner, F.G. and Dörner, F.
A novel vaccinia virus expression system allowing construction of
recombinants without the need for selection markers, plasmids and
bacterial hosts
J. Gen. Virol. 76 (Pt 12), 2957-2962 (1995)
JOURNAL
MEDLINE        96112180
AUTHORS        Pfleiderer, M.
TITLE          Direct Submission
SUBMITTED      (20-JUL-1995) M. Pfleiderer, Immuno AG, A. 2304
ORTH/DONAU, Uferstrasse 15, AUSTRIA
COMMENT        Related sequence M35027.
FEATURES       Location/Qualifiers
1..2164
/organism="Vaccinia virus"
/strain="WR"
/species="taxon:"mammalia"
/db_xref="taxon:10245"
/lab_host="mammalian cell lines"
/cell_line="recombinant Vaccinia virus"
/clone="VDPCL2"
/map="45478 nt number"
1..485
/note="Hind III F fragment"
479..1662
/note="NOT I gene cassette"
479
/note="NOT I cleavage site corresponding to the unique NOT
I site in the Vv genome at position 45478 (strain
copenhagen)"
486..543
/note="strong synthetic early/late promoter"
536
/note="unique Sfi I cleavage site"
544..594
/note="weak synthetic early/late promoter"
595..1620
/gene="hph"
595..1620
/gene="hph"
/codon_start=1
/product="hygromycin phosphotransferase"
/protein_id="CAA61953.1"
/db_xref="GI:953211"
/db_xref="SWISS-PROT:P00557"
/translation="MKKPELTATSVKFLIEKFDVSDLMOLSEGESRAFSFDVGG
GYVLRVNSCADGFYKDRYVYRHFASALPIPEVLDIGEFSESLTYCISRRAGVTLOD
LPETEPALVQPAEAMDAADLSQTSQGFPGFQGTWTRDTCIALADPHY
HWQTVMDTVASVAQALDELMLWEDCEVRLHVFADFGSNVLTIDNGRTAVLDWS
EAMFGDSQYEVANIFFRWPLACMEQOTYFFRRHPFELAGSPRLRAYMLRIGLDOLYQ
SLVDGNFDDAAWAQGRCDALVRSGAGTVGRTQIARRSAVWTDGCVLEADSGNRRP
TPRAKE"
1635
/note="Rsr II cleavage site"
1642
/note="NOT I cleavage site"
1643..2164
/note="Hind III F fragment"

Query Match          54.1%; Score 40.6; DB 16; Length 2164;
Best Local Similarity 73.2%; Pred. No. 0.038; 19; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 5 AAAAATTGAAATTTATTTATTTTGGGATATAAAGCGCGCCGCGCCGCGCGCCG 64
|||||
Db 488 AAAAATTGAAATTTATTTATTTTGGGATATAAAGCGCGCGCCGCGCCGCGCCATATG 547
|||||

QY 65 CCAACGCGCGGA 75
|||
Db 548 TAAAGTTGAA 558

RESULT 2
VH3FDPCL2/c
LOCUS          V.virus DNA for Hind III F fragment (clone VDPCL2).
DEFINITION
ACCESSION      X89857
VERSION        X89857.1 GI:927569
KEYWORDS       HindIII F fragment; hph gene; hygromycin phosphotransferase;
SOURCE         Vaccinia virus.
ORGANISM       Vaccinia virus
VIRUSES; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
Orthopoxvirus.
1 (bases 1 to 2164)
Pfleiderer, M., Falkner, F.G. and Dörner, F.
A novel vaccinia virus expression system allowing construction of
recombinants without the need for selection markers, plasmids and
bacterial hosts
J. Gen. Virol. 76 (Pt 12), 2957-2962 (1995)
JOURNAL
MEDLINE        96112180
AUTHORS        Pfleiderer, M.
TITLE          Direct Submission
SUBMITTED      (20-JUL-1995) M. Pfleiderer, Immuno AG, A. 2304
ORTH/DONAU, Uferstrasse 15, AUSTRIA
COMMENT        Related sequence M35027.
FEATURES       Location/Qualifiers
1..2164
/organism="Vaccinia virus"
/strain="WR"
/species="taxon:"mammalia"
/db_xref="taxon:10245"
/lab_host="mammalian cell lines"
/cell_line="recombinant Vaccinia virus"
/clone="VDPCL2"
/map="45478 nt number"
1..485
/note="Hind III F fragment"
479..1662
/note="NOT I gene cassette"
479
/note="NOT I cleavage site corresponding to the unique NOT
I site in the Vv genome at position 45478 (strain
copenhagen)"
486..543
/note="strong synthetic early/late promoter"
536
/note="unique Sfi I cleavage site"
544..594
/note="weak synthetic early/late promoter"
595..1620
/gene="hph"
595..1620
/gene="hph"
/codon_start=1
/product="hygromycin phosphotransferase"
/protein_id="CAA61953.1"
/db_xref="GI:953211"
/db_xref="SWISS-PROT:P00557"
/translation="MKKPELTATSVKFLIEKFDVSDLMOLSEGESRAFSFDVGG
GYVLRVNSCADGFYKDRYVYRHFASALPIPEVLDIGEFSESLTYCISRRAGVTLOD
LPETEPALVQPAEAMDAADLSQTSQGFPGFQGTWTRDTCIALADPHY
HWQTVMDTVASVAQALDELMLWEDCEVRLHVFADFGSNVLTIDNGRTAVLDWS
EAMFGDSQYEVANIFFRWPLACMEQOTYFFRRHPFELAGSPRLRAYMLRIGLDOLYQ
SLVDGNFDDAAWAQGRCDALVRSGAGTVGRTQIARRSAVWTDGCVLEADSGNRRP
TPRAKE"
1635
/note="Rsr II cleavage site"
1642
/note="NOT I cleavage site"
1643..2164
/note="Hind III F fragment"

```

LPETELPAVLQPVAEAMDAIAAADLSQTSFGPFGQIGQYTTWRDFICAIADPHVY
HWQIVMDTVSASVAQALDELMWAEDCEVVRHLVHADFGSNVLTNDNGRTAVIDWS
EAMFGDSQYEVANIFFRWPWACMEQOTYFERRHPELAGSPRLRAYMLRIGLDOLYO
SLVGNFDDAAWAQGRCDAIVRSGAGTVGRTOIARRSAAVWTDGCVVEVLADSGNRRPS
TRPRAKE"

promoter complement(1532..1585)
/note="weak synthetic early/late promoter"

misc_feature 1592
/note="sfi cleavage site"

promoter complement(1598..1639)
/note="strong synthetic early/late promoter"

misc_feature 1642
/note="NOT I cleavage site"

misc_feature 1643..2164
/note="Hind III F fragment"

BASE COUNT 595 a 524 c 439 g 606 t

ORIGIN

Query Match 54.1%; Score 40.6; DB 16; Length 2164;
Best Local Similarity 73.2%; Pred. No. 0.038; Mismatches 0; Gaps 0;
Matches 52; Conservative 0; Indels 0; Gaps 0;

QY 5 AAAAAATGAAATTTATTTTTTTTGGAAATATAAAGCGCGCCGATGGCGCGCGG 64

DB 1638 AAAAAATGAAATTTATTTTTTTTGGAAATATAAAGCGCGCGCCGATGGCGCGCGG 64

QY 65 CCAACGGCGGA 75

DB 1578 TAAAGTTGAA 1568

RESULT 3

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

IMMUNO AG (AT)

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

QY 3 CCAAAAATGAAATTTATTTTTTTTGGAAATATAAAGCGCGCGCCATGGGCC 58

DB 9 CTAAAAATGAAATTTATTTTTTTTGGAAATATAAAGCGCGCCATGGGCC 64

RESULT 4

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

A59040

Sequence 28 from Patent EP0753581.

A59040

A59040.1

GI:3714475

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

IMMUNO AG (AT)

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

QY 3 CCAAAAATGAAATTTATTTTTTTTGGAAATATAAAGCGCGCGCCATGGGCC 58

DB 80 CTAAAAATGAAATTTATTTTTTTTGGAAATATAAAGCGCGCCATGGGCC 25

RESULT 5

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

IMMUNO AG (AT)

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

QY 3 CCAAAAATGAAATTTATTTTTTTTGGAAATATAAAGCGCGCGCCATGGGCC 58

DB 9 CTAAAAATGAAATTTATTTTTTTTGGAAATATAAAGCGCGCCATGGGCC 64

RESULT 4

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

A59040

Sequence 28 from Patent EP0753581.

A59040

A59040.1

GI:3714475

unidentified.

unclassified.

unclassified.

1 (bases 1 to 86)

Scheiflinger,F.D., Antoine,G.D., Falkner, Falko-Guenter,D.,

Donner,F.P. and Eibl,J.D.

Improved recombinant eukaryotic cytoplasmic viruses, method for

their production and their use as vaccines

Patent: EP 0753581-A 28 15-JAN-1997;

IMMUNO AG (AT)

Location/Qualifiers

1..86

/organism="unidentified"

/db_xref="taxon:32644"

38 a 11 c 9 g 28 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

QY 3 CCAAAAATGAAATTTATTTTTTTTGGAAATATAAAGCGCGCGCCATGGGCC 58

DB 80 CTAAAAATGAAATTTATTTTTTTTGGAAATATAAAGCGCGCCATGGGCC 25

RESULT 5

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

IMMUNO AG (AT)

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

QY 3 CCAAAAATGAAATTTATTTTTTTTGGAAATATAAAGCGCGCGCCATGGGCC 58

DB 9 CTAAAAATGAAATTTATTTTTTTTGGAAATATAAAGCGCGCCATGGGCC 64

RESULT 4

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

A59040

Sequence 28 from Patent EP0753581.

Center: Genome Therapeutics Corporation

Center code: GTC

Web site: <http://www.genomecorp.com/>

Contact: gtc-seqcenter@genomecorp.com

Project information

Center project name: hg050

Summary Statistics

Sequencing vector: N/A

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 990315

Consensus quality: 130458 bases at least Q40

Consensus quality: 146152 bases at least Q30

Consensus quality: 149165 bases at least Q20

Insert size: 160821; sum-of-contigs

Quality coverage: 4.1x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 29 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1162: contig of 1162 bp in length
* 1163: contig of 1269 bp in length
* 2432: contig of 1436 bp in length
* 3868: contig of 1224 bp in length
* 5092: contig of 1861 bp in length
* 6953: contig of 1173 bp in length
* 8126: contig of 1393 bp in length
* 9519: contig of 1130 bp in length
* 10649: contig of 1111 bp in length
* 11760: contig of 1443 bp in length
* 13203: contig of 2231 bp in length
* 15434: contig of 2194 bp in length
* 17628: contig of 2029 bp in length
* 19657: contig of 3079 bp in length
* 22736: contig of 2384 bp in length
* 25120: contig of 2800 bp in length
* 27920: contig of 3730 bp in length
* 31650: contig of 5734 bp in length
* 37384: contig of 4217 bp in length
* 41601: contig of 3164 bp in length
* 44765: contig of 5890 bp in length
* 50655: contig of 7954 bp in length
* 58609: contig of 8037 bp in length
* 66646: contig of 9923 bp in length
* 76569: contig of 10040 bp in length
* 86609: contig of 14860 bp in length
* 101469: contig of 16559 bp in length
* 118028: contig of 22768 bp in length
* 140796: contig of 20026 bp in length.

FEATURES

source

1. .160821
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RP11-179H18"
/clone_lib="RP11-179H18"

43392 a 36669 c 36666 g 44015 t 79 others

BASE COUNT
ORIGIN

Query Match 43.2%; Score 32.4; DB 54; Length 160821;
Best Local Similarity 71.2%; Pred. No. 15;
Matches 42; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 6 AAAATTGAATTTTATTTTGTGAATATAAGCGCGCCCGCCATGGCGCGCGCG 64
Db 27681 ATAATTAACTTTTATTTTGTGAATATAAGCGCGCGCGCGCGCGCG 27739

RESULT 6

AC009246/c

LOCUS AC009246.7

DEFINITION Homo sapiens, *** SEQUENCING IN PROGRESS ***

AC009246

VERSION AC009246.7

KEYWORDS HTG; HTGS_PHASE1.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 331068)

Muzny,D.M., Adams,C., Aronson,A., Bailey,M., Barbaria,J.,

Blankenburg,K., Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C.,

Bunac,C., Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z.,

Cox,C., David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,

Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,

Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,

Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hoques,M.,

Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,

Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,

Lichtarge,O., Liu,J., Liu,W., Logan,O., Lu,J., Lucier,R.,

Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M., Morris,S.,

Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S., Osval,G.,

Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L., Quiles,M.,

Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S., Shah,E.,

Shen,H., Simon,M., Sparks,A., Stamps,A., Suggang,R., Tabor,P.,

Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M., Watlington,S.,

Weinstock,G., Weinstock,I.R., Williamson,A., Worley,K., Wren,J.,

Wrenford,G., Yu,W., Zhou,X. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 331068)

Worley,K.C.

Direct Submission

Submitted (08-AUG-1999) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Sep 2, 1999 this sequence version replaced gi:5787982.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 96 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 74449: contig of 74449 bp in length

* 74450: gap of unknown length

* 74470: contig of 16958 bp in length

* 91428: gap of unknown length

* 91448: contig of 15033 bp in length

* 106481: gap of unknown length

* 106501: gap of unknown length

* 119380: contig of 12879 bp in length

* 130144: gap of unknown length

* 130165: contig of 10745 bp in length

* 130165: gap of unknown length

* 140200: contig of 10035 bp in length

* 140220: gap of unknown length

* 149785: contig of 9545 bp in length

* 149785: gap of unknown length

* 157915: contig of 8130 bp in length

* 157935: gap of unknown length

* 164330: contig of 6395 bp in length

* 164350: gap of unknown length

* 170629: contig of 6260 bp in length

* 170630: gap of unknown length

[illegible]

*	1	855:	contig of 855 bp in length	*	25699	gap of unknown length	*
*			gap of unknown length	*		contig of 1102 bp in length	*
*	856	1440:	contig of 585 bp in length	*	26801	gap of unknown length	*
*			gap of unknown length	*		contig of 824 bp in length	*
*	1441	1718:	contig of 278 bp in length	*	27625	gap of unknown length	*
*			gap of unknown length	*		contig of 224 bp in length	*
*	1719	2205:	contig of 487 bp in length	*	27849	gap of unknown length	*
*			gap of unknown length	*		contig of 652 bp in length	*
*	2206	2703:	contig of 498 bp in length	*	28501	gap of unknown length	*
*			gap of unknown length	*		contig of 886 bp in length	*
*	2704	3509:	contig of 806 bp in length	*	29387	gap of unknown length	*
*			gap of unknown length	*		contig of 1097 bp in length	*
*	3510	4079:	contig of 570 bp in length	*	30484	gap of unknown length	*
*			gap of unknown length	*		contig of 249 bp in length	*
*	4080	4558:	contig of 479 bp in length	*	30733	gap of unknown length	*
*			gap of unknown length	*		contig of 1526 bp in length	*
*	4559	4812:	contig of 254 bp in length	*	32259	gap of unknown length	*
*			gap of unknown length	*		contig of 707 bp in length	*
*	4813	5923:	contig of 1111 bp in length	*	32966	gap of unknown length	*
*			gap of unknown length	*		contig of 574 bp in length	*
*	5924	6901:	contig of 978 bp in length	*	33540	gap of unknown length	*
*			gap of unknown length	*		contig of 1091 bp in length	*
*	6902	7515:	contig of 614 bp in length	*	34631	gap of unknown length	*
*			gap of unknown length	*		contig of 937 bp in length	*
*	7516	8535:	contig of 1020 bp in length	*	35568	gap of unknown length	*
*			gap of unknown length	*		contig of 97 bp in length	*
*	8536	8750:	contig of 215 bp in length	*	35665	gap of unknown length	*
*			gap of unknown length	*		contig of 669 bp in length	*
*	8751	9373:	contig of 623 bp in length	*	36334	gap of unknown length	*
*			gap of unknown length	*		contig of 383 bp in length	*
*	9374	10192:	contig of 819 bp in length	*	36717	gap of unknown length	*
*			gap of unknown length	*		contig of 401 bp in length	*
*	10193	10880:	contig of 688 bp in length	*	37118	gap of unknown length	*
*			gap of unknown length	*		contig of 1159 bp in length	*
*	10881	11050:	contig of 170 bp in length	*	38277	gap of unknown length	*
*			gap of unknown length	*		contig of 1006 bp in length	*
*	11051	11269:	contig of 219 bp in length	*	39283	gap of unknown length	*
*			gap of unknown length	*		contig of 371 bp in length	*
*	11270	12430:	contig of 1161 bp in length	*	39654	gap of unknown length	*
*			gap of unknown length	*		contig of 514 bp in length	*
*	12431	13216:	contig of 786 bp in length	*	40168	gap of unknown length	*
*			gap of unknown length	*		contig of 703 bp in length	*
*	13217	13642:	contig of 426 bp in length	*	40871	gap of unknown length	*
*			gap of unknown length	*		contig of 259 bp in length	*
*	13643	14688:	contig of 1046 bp in length	*	41129	gap of unknown length	*
*			gap of unknown length	*		contig of 270 bp in length	*
*	14689	14797:	contig of 109 bp in length	*	41399	gap of unknown length	*
*			gap of unknown length	*		contig of 254 bp in length	*
*	14798	15653:	contig of 856 bp in length	*	41654	gap of unknown length	*
*			gap of unknown length	*		contig of 637 bp in length	*
*	15654	16374:	contig of 721 bp in length	*	42290	gap of unknown length	*
*			gap of unknown length	*		contig of 1041 bp in length	*
*	16375	17396:	contig of 1022 bp in length	*	43331	gap of unknown length	*
*			gap of unknown length	*		contig of 1168 bp in length	*
*	17397	17504:	contig of 108 bp in length	*	44499	gap of unknown length	*
*			gap of unknown length	*		contig of 1196 bp in length	*
*	17505	18320:	contig of 816 bp in length	*	45695	gap of unknown length	*
*			gap of unknown length	*		contig of 812 bp in length	*
*	18321	19338:	contig of 1018 bp in length	*	46507	gap of unknown length	*
*			gap of unknown length	*		contig of 1015 bp in length	*
*	19339	20614:	contig of 1276 bp in length	*	47522	gap of unknown length	*
*			gap of unknown length	*		contig of 235 bp in length	*
*	20615	21816:	contig of 1202 bp in length	*	47757	gap of unknown length	*
*			gap of unknown length	*		contig of 244 bp in length	*
*	21817	22751:	contig of 935 bp in length	*	48001	gap of unknown length	*
*			gap of unknown length	*		contig of 610 bp in length	*
*	22752	23821:	contig of 1070 bp in length	*	48611	gap of unknown length	*
*			gap of unknown length	*		contig of 544 bp in length	*
*	23822	24608:	contig of 787 bp in length	*	49155	gap of unknown length	*
*			gap of unknown length	*		contig of 100 bp in length	*
*	24609	24769:	contig of 161 bp in length	*	49255	gap of unknown length	*
*			gap of unknown length	*		contig of 853 bp in length	*
*	24770	25698:	contig of 929 bp in length	*	50108	gap of unknown length	*


```

repeat_region 38460..38507
/rpt_family="(TG)n"
repeat_region 38618..38915
/rpt_family="AluX"
repeat_region complement(40279..40618)
/rpt_family="MER46B"
repeat_region complement(40768..41251)
/rpt_family="LiMA8"
repeat_region 41252..41284
/rpt_family="(TG)n"
repeat_region 41303..41388
/rpt_family="(TA)n"
repeat_region complement(41390..41557)
/rpt_family="LiMA7"
repeat_region 41558..41673
/rpt_family="(TA)n"

Query Match 40.3%; Score 30.2; DB 40; Length 161449;
Best Local Similarity 74.5%; Pred. No. 56;
Matches 38; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Oy 4 CAAAAATTCGAATTTATTTTTCGAAATAAAGCGGCCCATG 54
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101906 CAATACCTGAATCTATTTTTCGGAAGTATAGGTTCTCCATG 101856

RESULT 12
AC021672/c AC021672 161986 bp DNA HTG 19-JAN-2000
LOCUS Homo sapiens chromosome 11 clone RP11-88122 map 11, *** SEQUENCING
IN PROGRESS ***, 35 unordered pieces.
ACCESSION AC021672
VERSION AC021672.1 GI:6716080
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 161986)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 11, clone RP11-88122
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 161986)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Feenestor,J.,
Ferrelira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
Pierre,N., Pisanic,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
JOURNAL Submitted (19-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: LI363

```

Center clone name: 88_I_22

* NOTE: This is a 'working draft' sequence. It currently consists of 35 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1708: contig of 1708 bp in length
gap of unknown length
* 1709: contig of 1486 bp in length
gap of unknown length
* 3195: contig of 1469 bp in length
gap of unknown length
* 4664: contig of 1398 bp in length
gap of unknown length
* 6062: contig of 1729 bp in length
gap of unknown length
* 7791: contig of 1924 bp in length
gap of unknown length
* 9715: contig of 2456 bp in length
gap of unknown length
* 12171: contig of 2315 bp in length
gap of unknown length
* 14486: contig of 2595 bp in length
gap of unknown length
* 17081: contig of 2577 bp in length
gap of unknown length
* 19658: contig of 3272 bp in length
gap of unknown length
* 22930: contig of 2918 bp in length
gap of unknown length
* 25848: contig of 2147 bp in length
gap of unknown length
* 27995: contig of 3040 bp in length
gap of unknown length
* 31035: contig of 2733 bp in length
gap of unknown length
* 33758: contig of 2503 bp in length
gap of unknown length
* 36271: contig of 2931 bp in length
gap of unknown length
* 39202: contig of 3112 bp in length
gap of unknown length
* 42313: contig of 2427 bp in length
gap of unknown length
* 44740: contig of 3632 bp in length
gap of unknown length
* 48372: contig of 6062 bp in length
gap of unknown length
* 54434: contig of 5072 bp in length
gap of unknown length
* 59506: contig of 3905 bp in length
gap of unknown length
* 63411: contig of 3569 bp in length
gap of unknown length
* 66980: contig of 3975 bp in length
gap of unknown length
* 70955: contig of 4983 bp in length
gap of unknown length
* 75938: contig of 5427 bp in length
gap of unknown length
* 81365: contig of 6296 bp in length
gap of unknown length
* 87661: contig of 7946 bp in length
gap of unknown length
* 95607: contig of 8484 bp in length
gap of unknown length
* 104091: contig of 9948 bp in length
gap of unknown length
* 114039: contig of 9051 bp in length
gap of unknown length
* 123090: contig of 9051 bp in length
gap of unknown length

FEATURES	source	Location/Qualifiers
* 123091	132651:	gap of unknown length
* 132652	144718:	contig of 9561 bp in length
* 144719	161986:	gap of unknown length
* 144719	161986:	contig of 12067 bp in length
* 144719	161986:	gap of unknown length
* 144719	161986:	contig of 17268 bp in length.
BASE COUNT	45594 a 38630 c 35856 g 43893 t	13 others
ORIGIN		
FEATURES	source	Location/Qualifiers
1. .161986		gap of unknown length
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/chromosome="11"		
/map="11"		
/clone="RP11-88122"		
/clone_lib="RP11-11 Human Male BAC"		
45594 a 38630 c 35856 g 43893 t		13 others
Query Match	40.3%;	Score 30.2; DB 45; Length 161986;
Best Local Similarity	81.4%;	Pred. No. 66;
Matches	35; Conservative	0; Mismatches 8; Indels 0; Gaps 0;
QY	1	GGCCAAAATGAAATTTATTTTATTTTGGGAATATAAG 43
Db	34152	GGCCAAAAGTGAAGTTTTTTTTTTTTTTTTTTTTTTTAAAG 34110
RESULT	13	
AC009526	108061 bp	DNA PLN 11-DEC-1999
LOCUS	Arabidopsis thaliana chromosome I BAC F2J6	genomic sequence,
DEFINITION	complete sequence.	
AC009526	AC009526	
VERSION	AC009526.3	GI:6560894
KEYWORDS	HTG	
SOURCE	thale cress.	
ORGANISM	Arabidopsis thaliana	
REFERENCE	1 (bases 1 to 108061)	
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S., Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 108061)	
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S., Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.	
TITLE	Direct Submission	
JOURNAL	Submitted (26-AUG-1999)	DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	3 (bases 1 to 108061)	
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S., Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.	
TITLE	Direct Submission	
JOURNAL	Submitted (11-DEC-1999)	DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT	On Dec 11, 1999 this sequence version replaced gi:5881522. Bases 1-4149 of IGF clone F2J6 overlap with bases 91042-95190 of TBMU clone T10P12 (gb AC007203) and an unknown number of bases overlap at the end of F2J6 overlap with IGF clone F28H19 (gb AC006423).	

us-08-935-377-3.rge

Wed May 31 10:04:51 2000

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:58:06 ; Search time 1446.77 Seconds
(without alignments)
12.970 Million cell updates/sec

Title: US-08-935-377-3
Perfect score: 75
Sequence: 1 GCCCAAAATTTGAATTTA.....GCCCGCGCCACGCGGA 75

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	40	53.3	84	1	T78340	Chimeric virus con
2	40	53.3	86	1	T78341	Chimeric virus con
3	39.6	52.8	6474	1	Q40300	Sequence of plasm
4	38.8	51.7	5472	1	Q41005	Sequence of plasm
5	38.4	51.2	6811	1	Q41004	Sequence of plasm
6	38	50.7	42	1	Q67969	Sequence of modifi
7	38	50.7	51	1	Q40295	Sequence of oligo
8	38	50.7	53	1	Q40296	Sequence of oligo
9	38	50.7	3878	1	Q40299	Sequence of plasm
10	27	36.0	46	1	Q41030	pFP-Zsart construc
11	27	36.0	62	1	Q41025	P11 promoter mutan
12	27	36.0	88	1	Q40248	Sequence of plasm
13	27	36.0	92	1	Q40249	Sequence of plasm
14	26.6	35.5	97	1	Q40278	Sequence of synth
15	26.6	35.5	134	1	Q40250	Sequence of a DNA
16	26.6	35.5	4145	1	Q40279	Sequence of clone
17	26.6	35.5	25464	1	V57274	Human flavin-conta
18	26	34.7	2002	1	V36503	Human D2-2 gene as
19	25.8	34.4	90	1	V37789	Analytical solid p
20	25.8	34.4	4093	1	V58357	Coding sequence fo
21	25.6	34.1	42	1	Q41029	pFP-Zsart construc
22	25.4	33.9	126	1	V89886	EST clone CT317. N
23	25.4	33.9	3471	1	T17161	PSKH-1 cDNA. New i
24	25.2	33.6	55	1	Q41006	Sequence of modifi
25	25.2	33.6	232	1	T71192	Human genome fragm
26	25.2	33.6	3095	1	T31290	Mouse poly-immunog
27	25	33.3	1923	1	X51745	DNA encoding a hum
28	25	33.3	2453	1	V52449	Streptococcus pneu
29	25	33.3	3256	1	V29063	BRCA1 modulator pr
30	25	33.3	13425	1	V52284	Streptococcus pneu
31	25	33.3	16891	1	X37084	MEFV gene sequence
32	24.8	33.1	2436	1	V20438	Human c-N-ras onco
33	24.6	32.8	238	1	T26771	Human gene signatu

c 34 24.6 32.8 711 1 X00670 Human secreted pro
c 35 24.6 32.8 2058 1 V07560 Neocallimastix pat
c 36 24.6 32.8 3288 1 T05055 Human transforming
c 37 24.6 32.8 5040 1 N80916 Sequence encoding
c 38 24.4 32.5 91 1 Q41021 P11 promoter wildt
c 39 24.4 32.5 768 1 X39988 Prostate cancer as
c 40 24.4 32.5 812 1 X39990 Prostate cancer as
c 41 24.2 32.3 368 1 V87415 EST clone BR733. N
c 42 24.2 32.3 1296 1 V34289 Human secreted pro
c 43 24.2 32.3 5760 1 N50530 Sequence encoding
c 44 24.2 32.3 110000 1 V21209_13 Continuation (14 o
c 45 24 32.0 981 1 V60576 ERAB protein codin

ALIGNMENTS

RESULT 1
T78340
ID T78340 standard; DNA; 84 BP.
AC T78340;
DT 13-OCT-1997 (first entry)
DE Chimeric virus construction oligonucleotide oselpi.
KW Recombinant eukaryotic cytoplasmic DNA virus; vaccine production;
KW double selection marker; attenuated live pox virus; HIV; HBV;
KW surface antigen gene; encephalitis; tick-borne; ss.
OS Synthetic.
PN EP-753581-Al.
PD 15-JAN-1997.
PF 10-JUL-1995; 110727.
PR 10-JUL-1995; EP-110727.
PA (IMMO) IMMUNO AG.
PI Antoine G, Dornier F, Eibl J, Falkner F, Scheiflinger F;
DR WPI; 97-079382/08.
PT Prodn. of recombinant eukaryotic cytoplasmic DNA viruses for vaccine
PT prodn. - using novel construct contg. foreign DNA and double
PT selection marker
PS Example 9; Page 14; 68pp; English.
CC A method has been produced for the production of recombinant eukaryotic
CC cytoplasmic DNA viruses for vaccine production. The method involves
CC inserting a construct DNA molecule into the genome of a eukaryotic
CC cytoplasmic DNA virus, transfecting cells with the virus, and selecting
CC a recombinant virus that does not contain the double selection marker
CC cassette from the construct DNA. The DNA molecule preferably contains
CC one or more foreign DNA segments of a gene encoding an antigen of a
CC pathogen (especially HIV env, HIV gag, HIV gagpol, HIV nef, HBV
CC pRES1-S2-S surface antigen gene, HBV S1-S2-S surface antigen gene, HBV
CC S surface antigen gene, tick-borne encephalitis prem gene or tick-borne
CC encephalitis E-gene), and has a double selection marker cassette which
CC is flanked by at least two direct repeat DNA sequences, where at least
CC one foreign DNA segment does not occur within the DNA selection marker
CC by the direct repeat sequences and where the double selection marker
CC cassette comprises a dominant selection marker gene and a colour
CC selection marker gene (preferably the colour selection marker is the
CC Escherichia coli lacZ gene and the dominant selection marker is the
CC E.coli hph or gpt gene). The present sequence represents the oselpi
CC oligonucleotide used in the construction of chimeric MVA vaccinia and
CC fowlpox viruses expressing the tick-borne encephalitis virus prem and E
CC genes. The DNA constructs and method are used for the production of
CC vaccines, especially containing an attenuated live recombinant pox
CC virus.
SQ Sequence 84 BP; 28 A; 8 C; 10 G; 38 T;

Query Match 53.3%; Score 40; DB 1; Length 84;
Best Local Similarity 82.1%; Pred. No. 0.0012;
Matches 46; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 3 CCAAAATTTGAATTTATTTTTCGAATATATAAGCGCGCCATGGCC 58
Db 9 CTAAATTTGAATTTATTTTTCGAATATATAAGCGCGCCATGGCC 64

us-08-935-377-3.rng

Wed May 31 10:04:52 2000

```

RESULT 2
T78341/c
ID T78341 standard; DNA; 86 BP.
AC T78341;
DE 13-OCT-1997 (first entry)
DT Chimeric virus construction oligonucleotide osep2.
DI Recombinant eukaryotic cytoplasmic DNA virus; vaccine production;
KW double selection marker; attenuated live pox virus; HIV; HBV;
KW surface antigen gene; encephalitis; tick-borne; ss.
OS Synthetic.
PN EP-753581-A1.
PD 15-JAN-1997.
PF 10-JUL-1995; 110727.
PR 10-JUL-1995; EP-110727.
PA (IMMO ) IMMUNO AG.
PI Antoine G, Dorner F, Eibl J, Falkner F, Scheiflinger F;
DR WPI; 97-079382/08.
PT Prodn. of recombinant eukaryotic cytoplasmic DNA viruses for vaccine
FT prodn. - using novel construct contg. foreign DNA and double
FT selection marker
FT Example 9; Page 14; 68pp; English.
FS A method has been produced for the production of recombinant eukaryotic
CC cytoplasmic DNA viruses for vaccine production. The method involves
CC inserting a construct DNA molecule into the genome of a eukaryotic
CC cytoplasmic DNA virus, transfecting cells with the virus, and selecting
CC a recombinant virus that does not contain the double selection marker
CC cassette from the construct DNA. The DNA molecule preferably contains
CC one or more foreign DNA segments of a gene encoding an antigen of a
CC pathogen (especially HIV env, HIV gag, HIV gp120, HIV nef, HBV
CC p81-52-S surface antigen gene, HBV S1-S2-S surface antigen gene, HBV
CC S surface antigen gene, tick-borne encephalitis prem gene or tick-borne
CC encephalitis E gene), and has a double selection marker cassette which
CC is flanked by at least two direct repeat DNA sequences, where at least
CC one foreign DNA segment does not occur within the DNA sequence bounded
CC by the direct repeat sequences and where the double selection marker
CC cassette comprises a dominant selection marker gene and a colour
CC selection marker gene (preferably the colour selection marker is the
CC Escherichia coli lacZ gene and the dominant selection marker is the
CC E.coli hph or gpt gene). The present sequence represents the osep2
CC oligonucleotide used in the construction of chimeric MVA vaccinia and
CC fowlpox viruses expressing the tick-borne encephalitis virus prem and E
CC genes. The DNA constructs and method are used for the production of
CC vaccines, especially containing an attenuated live recombinant pox
CC virus.
SQ Sequence 86 BP; 38 A; 11 C; 9 G; 28 T;

Query Match 53.3%; Score 40; DB 1; Length 86;
Best Local Similarity 82.1%; Pred. No. 0.0012; Indels 0; Gaps 0;
Matches 46; Conservative 0; Mismatches 10;

QY 3 CCAAAATGAAATTTATTTATTTTGGAAATATAAGCGCGCCATGGGCC 58
DB 80 CTAATAATGAAATTTATTTATTTTGGAAATATAAGCGCCATGGGCC 25

RESULT 3
Q40300/c
ID Q40300 standard; DNA; 6474 BP.
AC Q40300;
DE 02-AUG-1993 (first entry)
DT Sequence of plasmid psep160MN.
DI Plasmid; cloning; restriction site; ss.
KW Synthetic.
OS Synthetic.
PN Location/Qualifiers
FT Key 1. .55
FT misc_feature /tag= a
FT /label= pTZ19R
FT complement (56. .108)
FT /tag= b
FT /label= Linker I in rc orientation
FT complement (110. .860)
FT /tag= c
FT cds

```

```

FT /label= E. coli gpt
FT complement (861. .1245)
FT /tag= d
FT /label= Vaccinia virus p7.5
FT /note = "Starting with the p7.5 internal NdeI site
FT at posn. 1241"
FT 1339. .1344
FT /tag= e
FT /label= HpaI site
FT complement (1259. .3916)
FT /tag= f
FT /label= HIV-MN env gene
FT complement (3917. .3970)
FT /tag= g
FT /label= vaccinia virus synthetic early/late
FT promoter
FT complement (3971. .4015)
FT /tag= h
FT /label= linker II
FT complement (4016. .6474)
FT /tag= i
FT /label= pTZ19R
FT AU9221269-A.
PN 04-WAR-1993.
PD 25-AUG-1992; 021269.
PF 26-AUG-1991; US-750080.
PR 20-JUL-1992; US-914738.
PA (IMMO ) IMMUNO AG.
PI Dorner F, Falkner FG, Pfeleiderer M, Scheiflinger F;
DR WPI; 91-126461/16.
PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
CC direct molecular cloning of modified DNA molecule contg.
CC cytoplasmic DNA virus genome
CC Example; Pages 169-172; 206pp; English.
PS The synthetic early/late promoter seip was used to express the
CC gp160-gene of the HIV-1 MN strain. For the construction of pL2
CC the 0.6kb XbaI-ClaI fragment of the plasmid pTM3 was substituted
CC by an XbaI-ClaI adaptor fragment consisting of the annealed
CC oligonucleotide o-542 and o-544. The intermediate plasmid
CC resulting from this cloning step was called pL1. The 0.84kb
CC AatII-SphI fragment were substituted by the AatII-SphI adaptor
CC fragment consisting of the annealed oligonucleotide o-541 and
CC o-543. The resulting plasmid was called pL2. The XbaI-SphI
CC fragment was treated with klenow-polymerase and inserted between
CC the PvuII sites of the plasmid pTZ19R. The resulting plasmid was
CC called pTZ-L2. The 0.6kb ClaI-NcoI fragment (the T7-promoter-EMC-
CC sequence) was replaced with a synthetic promoter fragment
CC consisting of the annealed oligonucleotide o-selpI and o-selpI.
CC The 239bp Sall-NdeI fragment of the resulting intermediate plasmid
CC was substituted by the Sall-NdeI adaptor consisting of the annealed
CC oligonucleotides o-830 and o-857. The resulting plasmid was called
CC psep1-gpt-L2. The 3.1kb env gene containing the EcoRI-PvuII
CC fragment of pMenvI was inserted into the EcoRI and StuI cut plasmid
CC psep1-gpt-L2 resulting in the intermediate plasmid psep1-gp160.1.
CC The 0.8kb NcoI-NsiI fragment of psep1-gp160 was substituted by a
CC PCR-generated 0.31kb NcoI-NsiI fragment resulting in the final
CC plasmid psep1-gp160MN. The primers used for the PCR reaction were
CC o-NcoI and o-NsiI.
SQ Sequence 6474 BP; 1648 A; 1532 C; 1410 G; 1880 T;

Query Match 52.8%; Score 39.6; DB 1; Length 6474;
Best Local Similarity 83.3%; Pred. No. 0.0036; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 9;

QY 5 AAAAATGAAATTTATTTATTTTGGAAATATAAGCGCGCCATGGGCC 58
DB 3964 AAAAATGAAATTTATTTATTTTGGAAATATAAGCGCCATGGGCC 3911

RESULT 4
Q41005/c
ID Q41005 standard; DNA; 5472 BP.

```


CC o-selPII. The 239bp Sall-NdeI fragment of the resulting
 CC intermediate plasmid was substituted by the Sall-NdeI adaptor
 CC consisting of the annealed oligonucleotides o-830 and o-857. The
 CC resulting plasmid was called pselp-gpt-L2.
 SQ Sequence 51 BP; 18 A; 4 C; 6 G; 23 T;

Query Match 50.7%; Score 38; DB 1; Length 51;
 Best Local Similarity 100.0%; Pred. No. 0.004;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AAAAATTGAAATTTATTTTATTTTGGGAATATAA 42
 DB 5 AAAAATTGAAATTTATTTTATTTTGGGAATATAA 42

RESULT 8
 Q40296/C
 ID Q40296 standard; DNA; 53 BP.
 AC Q40296;
 DT 02-AUG-1993 (first entry)
 DE Sequence of oligo o-selPII of a synthetic promoter fragment.
 KW Plasmid; cloning; restriction site; ss.
 OS Synthetic.
 PN AU9221269-A.
 PD 04-MAR-1993.
 PR 25-AUG-1992; 021269.
 PR 26-AUG-1991; US-750080.
 PR 20-JUL-1992; US-914738.
 PA (IMMO) IMMUNO AG.
 PI Dorner F, Falkner FG, Pfeleiderer M, Scheifflinger F;
 DR WPI; 93-126461/16.
 PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
 PT direct molecular cloning of modified DNA molecule contg.
 PT cytoplasmic DNA virus genome
 PS Example; Page 158; 206pp; English.
 CC The synthetic early/late promoter selp was used to express the
 CC gp160-gene of the HIV-1 MN strain. For the construction of pL2
 CC the 0.6kb XbaI-ClaI fragment of the plasmid pTM3 was substituted
 CC by an XbaI-ClaI adaptor fragment consisting of the annealed
 CC oligonucleotide o-542 and o-544. The intermediate plasmid
 CC resulting from this cloning step was called pL1. The 0.84kb
 CC AatII-SphI fragment were substituted by the AatII-SphI adaptor
 CC fragment consisting of the annealed oligonucleotide o-541 and
 CC o-543. The resulting plasmid was called pL2. The XbaI-SphI
 CC fragment was treated with Klenow-polymerase and inserted between
 CC the PvuII sites of the plasmid pT219R. The resulting plasmid was
 CC called pT2-L2. The 0.6kb ClaI-NcoI fragment (the
 CC T7-promoter-EMC-sequence) was replaced with a synthetic promoter
 CC fragment consisting of the annealed oligonucleotide o-selPI and
 CC o-selPII. The 239bp Sall-NdeI fragment of the resulting
 CC intermediate plasmid was substituted by the Sall-NdeI adaptor
 CC consisting of the annealed oligonucleotides o-830 and o-857. The
 CC resulting plasmid was called pselp-gpt-L2.
 SQ Sequence 53 BP; 24 A; 6 C; 4 G; 19 T;

Query Match 50.7%; Score 38; DB 1; Length 53;
 Best Local Similarity 100.0%; Pred. No. 0.004;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AAAAATTGAAATTTATTTTATTTTGGGAATATAA 42
 DB 51 AAAAATTGAAATTTATTTTATTTTGGGAATATAA 14

RESULT 9
 Q40299/C
 ID Q40299 standard; DNA; 3878 BP.
 AC Q40299;
 DT 02-AUG-1993 (first entry)
 DE Sequence of plasmid pselp-gpt-L2.
 KW Plasmid; cloning; restriction site; ss.

KW prevention; therapy; epitope; hepatitis B virus; ss.
 OS Synthetic.
 PN W09412617-A.
 PD 09-JUN-1994.
 PF 24-NOV-1993; U11474.
 PR 25-NOV-1992; US-982211.
 PA (ITBI-) INT BIOTECHNOLOGY LAB INC.
 PI Bernstine EG, Lewis T, Okeefe RW, Souw PTS;
 DR WPI; 94-200247/24.
 PT Prevention and treatment of hepatitis - using recombinant
 PT replicable vaccinia viruses contg. hepatitis B virus surface and
 PT core antigen nucleotide sequences
 PS Disclosure; Fig 6; 252pp; English.
 CC HBV core antigen (Ag) encoding sequences were subcloned and
 CC engineered so as to be transcriptionally controlled by a
 CC vaccinia or vaccinia-like promoter. A strong vaccinia promoter
 CC p7.5 (cochran et al., 1985, J. Virol. 54:30-37), having the
 CC sequence in Q67967, can be one of several vaccinia or vaccinia-
 CC like promoter used to direct expression of the HBV sequences.
 CC the p7.5 promoter is so named because it normally directs
 CC expression of vaccinia polypeptide of 7.5 kD. Another vaccinia
 CC promoter, p11, in nature directs expression of vaccinia structural
 CC protein of 11 kD. Unlike p7.5, p11 is active only late in viral
 CC replication. Its sequence is given in Q67968. The modified p7.5
 CC promoter (Q67969) is a strong synthetic promoter that is active
 CC in both early and late viral replication. Its sequence is partly
 CC based on the p7.5 promoter.
 SQ Sequence 42 BP; 17 A; 0 C; 4 G; 21 T;

Query Match 50.7%; Score 38; DB 1; Length 42;
 Best Local Similarity 100.0%; Pred. No. 0.0038;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AAAAATTGAAATTTATTTTATTTTGGGAATATAA 42
 DB 1 AAAAATTGAAATTTATTTTATTTTGGGAATATAA 38

RESULT 7
 Q40295
 ID Q40295 standard; DNA; 51 BP.
 AC Q40295;
 DT 02-AUG-1993 (first entry)
 DE Sequence of oligo o-selPI of a synthetic promoter fragment.
 KW Plasmid; cloning; restriction site; ss.
 OS Synthetic.
 PN AU9221269-A.
 PD 04-MAR-1993.
 PR 25-AUG-1992; 021269.
 PR 26-AUG-1991; US-750080.
 PR 20-JUL-1992; US-914738.
 PA (IMMO) IMMUNO AG.
 PI Dorner F, Falkner FG, Pfeleiderer M, Scheifflinger F;
 DR WPI; 93-126461/16.
 PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
 PT direct molecular cloning of modified DNA molecule contg.
 PT cytoplasmic DNA virus genome
 PS Example; Page 158; 206pp; English.
 CC The synthetic early/late promoter selp was used to express the
 CC gp160-gene of the HIV-1 MN strain. For the construction of pL2
 CC by an XbaI-ClaI fragment of the plasmid pTM3 was substituted
 CC the 0.6kb XbaI-ClaI adaptor fragment consisting of the annealed
 CC oligonucleotide o-542 and o-544. The intermediate plasmid
 CC resulting from this cloning step was called pL1. The 0.84kb
 CC AatII-SphI fragment were substituted by the AatII-SphI adaptor
 CC fragment consisting of the annealed oligonucleotide o-541 and
 CC o-543. The resulting plasmid was called pL2. The XbaI-SphI
 CC fragment was treated with Klenow-polymerase and inserted between
 CC the PvuII sites of the plasmid pT219R. The resulting plasmid was
 CC called pT2-L2. The 0.6kb ClaI-NcoI fragment (the
 CC T7-promoter-EMC-sequence) was replaced with a synthetic promoter
 CC fragment consisting of the annealed oligonucleotide o-selPI and

```

OS Synthetic.
FH key Location/Qualifiers
FT misc_feature 1. .55
FT FT /tag= a
FT FT /label= p7z19r
FT FT complement (56. .108)
FT FT /tag= b
FT FT /label= Linker I in rc orientation
FT FT complement (110. .860)
FT FT /tag= c
FT FT /label= E. coli gpt
FT FT complement (861. .1245)
FT FT /tag= d
FT FT /label= Vaccinia virus p7.5
FT FT /note= "starting with the p7.5 internal NdeI site
FT FT at posn. 1241"
FT FT 1339. .1344
FT FT /tag= e
FT FT /label= HpaI site
FT FT complement (1259. .1322)
FT FT /tag= g
FT FT /label= multiple cloning site
FT FT complement (1323. .1374)
FT FT /tag= i
FT FT /label= vaccinia virus synthetic early/late
FT FT promoter
FT FT complement (1375. .1414)
FT FT /tag= j
FT FT /label= linker II
FT FT complement (1415. .3878)
FT FT /tag= k
FT FT /label= p7z19r
FT FT
PN A09221269-A.
PD 04-MAR-1993.
PF 25-AUG-1992. 021269.
PR 26-AUG-1991. US-750080.
PR 20-JUL-1992. US-914738.
PA (IMMO ) IMMUNO AG.
PI Dorner F, Falkner FG, Pfeleiderer M, Scheifflinger F;
DR WPI; 93-126461/16.
PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT direct molecular cloning of modified DNA molecule contg.
PT cytoplasmic DNA virus genome
PS Example; Pages 165-167; 206pp; English.
CC The synthetic early/late promoter seip was used to express the
CC gp160-gene of the HIV-1 MN strain. For the construction of pL2
CC the 0.6kb XbaI-ClaI fragment of the plasmid pTM3 was substituted
CC by an XbaI-ClaI adaptor fragment consisting of the annealed
CC oligonucleotide o-342 and o-544. The intermediate plasmid
CC resulting from this cloning step was called pL1. The 0.84kb
CC AatII-SphI fragment were substituted by the AatII-SphI adaptor
CC fragment consisting of the annealed oligonucleotide o-541 and
CC o-543. The resulting plasmid was called pL2. The XbaI-SphI
CC fragment was treated with Klenow-polymerase and inserted between
CC the PvuII sites of the plasmid p7z19r. The resulting plasmid was
CC called p7z-L2. The 0.6kb ClaI-NcoI fragment (the 17-promoter-EMC-
CC sequence) was replaced with a synthetic promoter fragment
CC consisting of the annealed oligonucleotide o-selpl and o-selplI.
CC The 239bp SalI-NdeI fragment of the resulting intermediate plasmid
CC was substituted by the SalI-NdeI adaptor consisting of the annealed
CC oligonucleotides o-830 and o-857. The resulting plasmid was called
CC pseip-gpt-L2.
SQ Sequence 3878 BP; 1012 A; 927 C; 974 G; 965 T;

Query Match 50.78; Score 38; DB 1; Length 3878;
Best Local Similarity 100.0%; Pred. No. 0.0099;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AAAAATTGAATTTATTTTTTTTGGGAATATAA 42
Db 1368 AAAAATTGAATTTATTTTTTTTGGGAATATAA 1331

OS Synthetic.
FH key Location/Qualifiers
FT misc_feature 1. .55
FT FT /tag= a
FT FT /label= p7z19r
FT FT complement (56. .108)
FT FT /tag= b
FT FT /label= Linker I in rc orientation
FT FT complement (110. .860)
FT FT /tag= c
FT FT /label= E. coli gpt
FT FT complement (861. .1245)
FT FT /tag= d
FT FT /label= Vaccinia virus p7.5
FT FT /note= "starting with the p7.5 internal NdeI site
FT FT at posn. 1241"
FT FT 1339. .1344
FT FT /tag= e
FT FT /label= HpaI site
FT FT complement (1259. .1322)
FT FT /tag= g
FT FT /label= multiple cloning site
FT FT complement (1323. .1374)
FT FT /tag= i
FT FT /label= vaccinia virus synthetic early/late
FT FT promoter
FT FT complement (1375. .1414)
FT FT /tag= j
FT FT /label= linker II
FT FT complement (1415. .3878)
FT FT /tag= k
FT FT /label= p7z19r
FT FT
PN A09221269-A.
PD 04-MAR-1993.
PF 25-AUG-1992. 021269.
PR 26-AUG-1991. US-750080.
PR 20-JUL-1992. US-914738.
PA (IMMO ) IMMUNO AG.
PI Dorner F, Falkner FG, Pfeleiderer M, Scheifflinger F;
DR WPI; 93-126461/16.
PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT direct molecular cloning of modified DNA molecule contg.
PT cytoplasmic DNA virus genome
PS Example; Pages 165-167; 206pp; English.
CC The synthetic early/late promoter seip was used to express the
CC gp160-gene of the HIV-1 MN strain. For the construction of pL2
CC the 0.6kb XbaI-ClaI fragment of the plasmid pTM3 was substituted
CC by an XbaI-ClaI adaptor fragment consisting of the annealed
CC oligonucleotide o-342 and o-544. The intermediate plasmid
CC resulting from this cloning step was called pL1. The 0.84kb
CC AatII-SphI fragment were substituted by the AatII-SphI adaptor
CC fragment consisting of the annealed oligonucleotide o-541 and
CC o-543. The resulting plasmid was called pL2. The XbaI-SphI
CC fragment was treated with Klenow-polymerase and inserted between
CC the PvuII sites of the plasmid p7z19r. The resulting plasmid was
CC called p7z-L2. The 0.6kb ClaI-NcoI fragment (the 17-promoter-EMC-
CC sequence) was replaced with a synthetic promoter fragment
CC consisting of the annealed oligonucleotide o-selpl and o-selplI.
CC The 239bp SalI-NdeI fragment of the resulting intermediate plasmid
CC was substituted by the SalI-NdeI adaptor consisting of the annealed
CC oligonucleotides o-830 and o-857. The resulting plasmid was called
CC pseip-gpt-L2.
SQ Sequence 3878 BP; 1012 A; 927 C; 974 G; 965 T;

Query Match 50.78; Score 38; DB 1; Length 3878;
Best Local Similarity 100.0%; Pred. No. 0.0099;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AAAAATTGAATTTATTTTTTTTGGGAATATAA 42
Db 1368 AAAAATTGAATTTATTTTTTTTGGGAATATAA 1331

```

```

RESULT 10
Q41030 Q41030 standard; DNA; 46 BP.
ID Q41030;
AC Q41030;
DE 29-JUL-1993 (first entry)
DE PFP-tsart construction oligomer #2.
KW Fowl pox virus; FPV; P2; promoter; vaccinia virus; VV; mutation; apoAI;
KW multiple cloning site; plasmid; thymidine kinase; tk; Marck's disease;
KW intergenic region; post-translational modification; factor II; apoAII;
KW factor V; factor VII; factor VIII; factor IX; tick-borne encephalitis;
KW factor X; factor XI; factor XII; factor XIII; protein C;
KW hepatitis E; factor X; factor XI; factor XII; factor XIII; protein C;
KW plasminogen; apolipo protein; viral antigen; infectious bronchitis;
KW TBE; tetanus; HIV; HSV; pertussis; Newcastle disease; vaccine; ss.
OS Synthetic.
PN A09219580-A.
PD 04-MAR-1993.
PF 10-JUL-1992; 019580.
PR 26-AUG-1991; EP-114300.
PA (IMMO ) IMMUNO AG.
PI Dorner F, Falkner FG, Scheifflinger F;
DR WPI; 93-126452/16.
PT Prepn. of recombinant fowl pox virus for e.g. expression of
PT proteins - by inserting foreign DNA into enlarged inter-genic
PT region so FPV tk-gene remains intact and codes for entire
PT thymidine kinase
PT Disclosure; Page 39; 92pp; English.
CC The sequences given in Q41027-45 are oligomers which were used in the
CC construction of the insertion plasmids of the invention. These
CC plasmids comprised the fowl pox virus (FPV) P2 promoter sequence (see
CC Q41013 and Q41018-20), multiple cloning sites (see Q41015-16), the FPV
CC thymidine kinase (tk) gene and the tk gene intergenic region (see also
CC Q41011-12). The P2 promoter belongs to one of the strongest natural
CC promoters in vaccinia virus (VV) infected cells. These plasmids may be
CC used to express foreign proteins which require post-translational
CC modification eg. Factors II, V, VII, VIII, IX, X, XI, XII, XIII,
CC proteins C and S, von Willebrand factor, plasminogen and derivatives,
CC apolipo proteins, eg. apoAI and apoAII, and viral antigens, eg.
CC hepatitis B- antigens and the antigens of hepatitis C and E viruses,
CC of tickborne encephalitis (TBE) virus, antigens of HIV, HSV and whole
CC or partial sequences of antigens which cause pertussis, infectious
CC bronchitis, tetanus, malaria, Marck's disease and Newcastle disease,
CC these antigens being useful as vaccines.
SQ Sequence 46 BP; 9 A; 4 C; 9 G; 24 T;

Query Match 36.08; Score 27; DB 1; Length 46;
Best Local Similarity 85.7%; Pred. No. 6.5;
Matches 30; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 16 TTTTATTTTTTTTTTGGGAATATAAAGCGGCGCG 50
Db 11 TTTTATTTTTTTTTTGGGAATATAAATAGGCTGC 45

RESULT 11
Q41025 Q41025 standard; DNA; 62 BP.
ID Q41025;
AC Q41025;
DE 29-JUL-1993 (first entry)
DE P11 promoter mutant sart.
KW Fowl pox virus; FPV; P11; promoter; vaccinia virus; mutation; apoAI;
KW multiple cloning site; plasmid; thymidine kinase; tk; Marck's disease;
KW intergenic region; post-translational modification; factor II; apoAII;
KW factor V; factor VII; factor VIII; factor IX; tick-borne encephalitis;
KW factor X; factor XI; factor XII; factor XIII; protein C; VV;
KW von Willebrand factor; hepatitis B; hepatitis C; hepatitis E; malaria;
KW plasminogen; apolipo protein; viral antigen; infectious bronchitis;
KW TBE; tetanus; HIV; HSV; pertussis; Newcastle disease; vaccine; ss.
OS Synthetic.
PN Key Location/Qualifiers
FH Key 34. .38
FT promoter

```

Wed May 31 10:04:52 2000

FT /*tag= a
 FT /note= "Vaccinia virus late promoter consensus
 FT sequence"
 PN AU9219380-A.
 PD 04-MAR-1993.
 PF 10-JUL-1992; 019580.
 PR 26-AUG-1991; EP-114300.
 PA (IMMO) IMMUNO AG.
 PI Dörner F, Falkner FG, Scheifflinger F;
 DR WPI: 93-126452/16.
 PT Prep. of recombinant fowl pox virus for e.g. expression of
 PT proteins - by inserting foreign DNA into enlarged inter-genic
 PT region so FPV TK-gene remains intact and codes for entire
 PT thymidine kinase
 PS Disclosure; Fig 18; 92pp; English.
 CC The sequences given in Q41021-25 represent mutations of the vaccinia
 CC virus (VV) p11 promoter sequence which were included in the expression
 CC plasmids of the invention. The sequences were unclear when given in
 CC the specification and so represent the "best guess" of sequence. These
 CC plasmids further comprised multiple cloning sites (see Q41015-16), the
 CC FPV thymidine kinase (tk) gene and the tk gene intergenic region (see
 CC also Q41011-12). These plasmids may be used to express foreign
 CC proteins which require post-translational modification eg. Factors II,
 CC V, VII, VIII, IX, X, XI, XII, XIII, proteins C and S, von Willebrand
 CC factor, plasminogen and derivatives, apolipo proteins, eg. apoAI and
 CC apoAII, and viral antigens, eg. hepatitis B- antigens and the antigens
 CC of hepatitis C and E viruses, of tick-borne encephalitis (TBE) virus,
 CC antigens of HIV, HSV and whole or partial sequences of antigens which
 CC cause pertussis, Infectious bronchitis, tetanus, malaria, Marck's
 CC disease and Newcastle disease, these antigens being useful as vaccines.
 CC Sequence 62 BP; 14 A; 4 C; 11 G; 31 T;

Query Match 36.0%; Score 27; DB 1; Length 62;
 Best Local Similarity 85.7%; Pred. No. 6.9;
 Matches 30; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 16 TTTTATTTTTTTTTTGGGAATATAAAGCGCGCGC 50
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 11 TTTTATTTTTTTTTTGGCATATAAATAGGCTGC 45

RESULT 12
 Q40248 ID Q40248 standard; DNA; 88 BP.
 AC Q40248;
 DT 02-AUG-1993 (first entry)
 DE Sequence of plasmid pAI-S2 which comprises a first master cloning
 DE site comprised of unique sites of a second master cloning site of
 DE the vaccinia virus vector designated vdrk with a poxvirus promoter
 DE operatively linked to the first master cloning site.
 KW Plasmid; cloning; restriction site; ss.
 OS Synthetic.
 PN AU9221269-A.
 PD 04-MAR-1993.
 PF 25-AUG-1992; 021269.
 PR 26-AUG-1991; US-750080.
 PR 20-JUL-1992; US-914738.
 PA (IMMO) IMMUNO AG.
 PI Dörner F, Falkner FG, Pfeleiderer M, Scheifflinger F;
 DR WPI: 93-126461/16.
 PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
 PT direct molecular cloning of modified DNA molecule contg.
 PT cytoplasmic DNA virus genome
 PS Claim 63; Page 140; 206pp; English.
 CC Plasmids pAI-S2 and pA2-S2 comprise gene expression cassettes
 CC suitable for association of open reading frames already having a
 CC translation start codon with a synthetic poxvirus promoter (S2),
 CC prior to direct molecular transfer into vaccinia virus vector vdrk
 CC by forced cloning. The S2 promoter is present in different
 CC orientations in the two plasmids.
 CC Sequence 88 BP; 19 A; 13 C; 20 G; 36 T;

Query Match 36.0%; Score 27; DB 1; Length 88;
 Best Local Similarity 85.7%; Pred. No. 7.4;
 Matches 30; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 16 TTTTATTTTTTTTTTGGGAATATAAAGCGCGCGC 50
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 31 TTTTATTTTTTTTTTGGCATATAAATAGGCTGC 65

RESULT 13
 Q40249/c ID Q40249 standard; DNA; 92 BP.
 AC Q40249;
 DT 02-AUG-1993 (first entry)
 DE Sequence of plasmid pA2-S2 which comprises a first master cloning
 DE site comprised of unique sites of a second master cloning site of
 DE the vaccinia virus vector designated vdrk with a poxvirus promoter
 DE operatively linked to the first master cloning site.
 KW Plasmid; cloning; restriction site; ss.
 OS Synthetic.
 PN AU9221269-A.
 PD 04-MAR-1993.
 PF 25-AUG-1992; 021269.
 PR 26-AUG-1991; US-750080.
 PR 20-JUL-1992; US-914738.
 PA (IMMO) IMMUNO AG.
 PI Dörner F, Falkner FG, Pfeleiderer M, Scheifflinger F;
 DR WPI: 93-126461/16.
 PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
 PT direct molecular cloning of modified DNA molecule contg.
 PT cytoplasmic DNA virus genome
 PS Claim 63; Page 140; 206pp; English.
 CC Plasmids pAI-S2 and pA2-S2 comprise gene expression cassettes
 CC suitable for association of open reading frames already having a
 CC translation start codon with a synthetic poxvirus promoter (S2),
 CC prior to direct molecular transfer into vaccinia virus vector vdrk
 CC by forced cloning. The S2 promoter is present in different
 CC orientations in the two plasmids.
 CC Sequence 92 BP; 37 A; 20 C; 14 G; 21 T;

Query Match 36.0%; Score 27; DB 1; Length 92;
 Best Local Similarity 85.7%; Pred. No. 7.5;
 Matches 30; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 16 TTTTATTTTTTTTTTGGGAATATAAAGCGCGCGC 50
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 62 TTTTATTTTTTTTTTGGCATATAAATAGGCTGC 28

RESULT 14
 Q40278/c ID Q40278 standard; DNA; 97 BP.
 AC Q40278;
 DT 02-AUG-1993 (first entry)
 DE Sequence of synthetic linker consisting of the oligonucleotide
 DE P-artp(9) abd P-artp(10) used to construct pN2gpt-S4.
 KW Plasmid; cloning; restriction site; ss.
 OS Synthetic.
 PN AU9221269-A.
 PD 04-MAR-1993.
 PF 25-AUG-1992; 021269.
 PR 26-AUG-1991; US-750080.
 PR 20-JUL-1992; US-914738.
 PA (IMMO) IMMUNO AG.
 PI Dörner F, Falkner FG, Pfeleiderer M, Scheifflinger F;
 DR WPI: 93-126461/16.
 PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
 PT direct molecular cloning of modified DNA molecule contg.
 PT cytoplasmic DNA virus genome
 PS Example; Page 153; 206pp; English.
 CC Plasmids pN2gpt-S3A and pN2gpt-S4 comprise expression cassettes

CC with a selective marker. These plasmids were constructed by first
CC making plasmids pN2-gpta and pN2-gptb which contain an E. coli gpt
CC gene driven by the vaccinia virus P7.5 promoter, flanked by
CC several unique restriction sites including NotI. For the
CC construction of plasmid pN2gpt-S3A, the parental plasmid pN2-gptb
CC was digested with PstI and ClaI and ligated with a synthetic
CC linker consisting of the oligonucleotides P-artP(7) and P-artP(8).
CC The synthetic promoter sequence of pN2gpt-S3A corresp. to the
CC oligonucleotide P-artP(7). For the construction of plasmid
CC pN2gpt-S4, the plasmid pN2-gptb was digested with PstI and ClaI
CC and ligated with an adaptor sequence consisting of the
CC oligonucleotides P-artP(9) and P-artP(10). The synthetic promoter
CC sequence of pN2gpt-S4 corresp. to the oligonucleotide P-artP(9).
SQ Sequence 97 BP; 32 A; 23 C; 22 G; 20 T;

Query Match 35.5%; Score 26.6; DB 1; Length 97;
Best Local Similarity 66.7%; Pred. No. 9.9; Indels 0; Gaps 0;
Matches 38; Conservative 0; Mismatches 19;

QY 16 TTTTATTTTTTTTGAATATAAGCGCGCCGCGCATGGCGCCGCCCAACGGC 72
||||| ||||||| ||||||| || ||| |||||
Db 85 TTTTATTTTTTTTGGCATATAAATCGTTAACGAATTCATGCGCCGGGAAGGC 29

RESULT 15

Q40250
ID Q40250 standard; DNA; 134 BP.
AC Q40250;
DT 02-AUG-1993 (first entry)
DE Sequence of a DNA segment comprising a sequence-specific
DE endonuclease cleavage site with a cleavage site for the bacterial
DE restriction endonuclease NotI at each end and a selective marker
DE gene and a second poxvirus promoter in plasmid pN2-gpt-S4.
KW Plasmid; cloning; restriction site; ss.
OS Synthetic.
PN AU9221269-A.
PD 04-MAR-1993.
PF 25-AUG-1992; 021269.
PR 26-AUG-1991; US-750080.
PR 20-JUL-1992; US-914738.
PA (IMMO) IMMUNO AG.
PI Dörner F, Falkner FG, Pfeleiderer M, Scheifflinger F;
DR WPI; 93-126461/16.
PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT direct molecular cloning of modified DNA molecule contg.
PT cytoplasmic DNA virus genome
PS Claim 58; Page 141; 206pp; English.
CC Plasmid pN2 contains a DNA segment having a sequence-specific
CC endonuclease cleavage site that is unique in the plasmid with a NotI
CC site at each end. In this plasmid the DNA segment may further
CC comprise a selective marker gene under transcriptional control of a
CC poxvirus promoter, for instance pN2-gpta and pN2-gptb. The DNA
CC segment may further comprise a second poxvirus promoter operatively
CC linked to a DNA sequence comprising a restriction endonuclease
CC cleavage site, e.g. pN2gpt-S4. pN2gpt-S4 are comprised of gene
CC expression cassettes suitable for association of an open reading
CC frame having a translation start codon, with a synthetic
CC promoter S4 prior to direct molecular transfer into a unique site in
CC vaccinia virus vecotr vATK.
SQ Sequence 134 BP; 26 A; 34 C; 35 G; 39 T;

Query Match 35.5%; Score 26.6; DB 1; Length 134;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 38; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 16 TTTTATTTTTTTTGAATATAAGCGCGCCGCGCATGGCGCCGCCCAACGGC 72
||||| ||||||| ||||||| || ||| |||||
Db 29 TTTTATTTTTTTTGGCATATAAATCGTTAACGAATTCATGCGCCGGGAAGGC 85

Search completed: May 29, 2000, 21:58:11
Job time: 38159 sec

us-08-935-377-3.rng

Wed May 31 10:04:52 2000

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 22:08:49 ; Search time 621.83 Seconds
(without alignments)
15.678 Million cell updates/sec

Title: US-08-935-377-3

Perfect score: 75
Sequence: 1 GGCACAAATGAAATTTA.....GCCCGGCCGCCACGGCGGA 75

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgnl_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgnl_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgnl_6/ptodata/1/ina/5C_COMB.seq.*
4: /cgnl_6/ptodata/1/ina/5D_COMB.seq.*
5: /cgnl_6/ptodata/1/ina/5E_COMB.seq.*
6: /cgnl_6/ptodata/1/ina/5F_COMB.seq.*
7: /cgnl_6/ptodata/1/ina/5G_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	36.0	46	1	US-08-232-463-40
2	27	36.0	63	1	US-08-232-463-47
3	27	36.0	88	1	US-07-750-080A-11
4	27	36.0	92	1	US-07-750-080A-12
5	27	36.0	9890	1	US-08-232-463-18
6	26.6	35.5	97	1	US-07-750-080A-41
7	26.6	35.5	134	1	US-07-750-080A-14
8	26	34.7	1000	3	US-08-747-121-20
9	26	34.7	2002	3	US-08-747-121-1
10	25.4	33.9	3471	3	US-08-715-568A-2
11	25.2	33.6	55	1	US-07-750-080A-38
12	25.2	33.6	55	1	US-07-750-080A-39
13	25.2	33.6	3095	5	US-08-434-000A-7
14	25	33.3	3256	4	US-08-968-751-3
15	24.8	33.1	2436	2	US-08-306-691B-16
16	24.6	32.8	834	3	US-08-967-101-113
17	24.6	32.8	834	4	US-08-592-541-113
18	24.6	32.8	2058	4	US-08-749-391-1
19	24.6	32.8	3288	1	US-08-208-008C-3
20	23.8	31.7	42	4	US-08-484-575A-11
21	23.8	31.7	42	5	US-08-477-459-11
22	23.8	31.7	42	6	PCT-US94-01826A-11
23	23.8	31.7	42	6	PCT-US94-02252A-11
24	23.8	31.7	49	6	PCT-US96-07709-39
25	23.8	31.7	49	6	PCT-US96-07709-40
26	23.8	31.7	49	6	PCT-US96-09848-19

ALIGNMENTS

RESULT 1
US-08-232-463-40
; Sequence 40, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: sart(4)
US-08-232-463-40

c 27 23.8 31.7 49 6 PCT-US96-09848-20
c 28 23.8 31.7 55 1 US-07-803-633A-3
c 29 23.8 31.7 55 1 US-07-803-633A-4
c 30 23.8 31.7 55 3 US-08-525-742-13
c 31 23.8 31.7 55 3 US-08-525-742-14
c 32 23.8 31.7 70 3 US-08-097-554A-43
c 33 23.8 31.7 70 5 US-08-480-640A-43
c 34 23.8 31.7 74 3 US-08-097-554A-44
c 35 23.8 31.7 74 5 US-08-480-640A-44
c 36 23.8 31.7 91 3 US-08-525-742-51
c 37 23.8 31.7 93 1 US-07-750-080A-40
c 38 23.8 31.7 93 3 US-08-525-742-28
c 39 23.8 31.7 95 3 US-08-525-742-29
c 40 23.8 31.7 96 3 US-08-525-742-30
c 41 23.8 31.7 97 3 US-08-525-742-27
c 42 23.8 31.7 102 1 US-07-820-154A-10
c 43 23.8 31.7 102 3 US-08-097-554A-10
c 44 23.8 31.7 102 5 US-08-480-640A-10
c 45 23.8 31.7 102 6 PCT-US93-00324-10
Sequence 20, Appl
Sequence 3, Appl
Sequence 4, Appl
Sequence 13, Appl
Sequence 14, Appl
Sequence 43, Appl
Sequence 43, Appl
Sequence 44, Appl
Sequence 44, Appl
Sequence 51, Appl
Sequence 40, Appl
Sequence 28, Appl
Sequence 29, Appl
Sequence 30, Appl
Sequence 27, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 10, Appl

us-08-935-377-3.rni

Wed May 31 10:04:52 2000

```
Query Match          36.0%; Score 27; DB 1; Length 46;
Best Local Similarity 85.7%; Pred. No. 1.6;
Matches 30; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 16 TTTTATTTTTTTTTTGGGAATATAAAGCGCGC 50
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11 TTTTATTTTTTTTTTGGCATATAAATAGGCTGC 45

RESULT 2
US-08-232-463-47
; Sequence 47, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF A MODIFIED
; TITLE OF INVENTION: EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07750,080A
; FILING DATE: 19910826
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/106 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 88 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: PAL-S2 (Fig. 4.5)
; US-07-750-080A-11

Query Match          36.0%; Score 27; DB 1; Length 88;
Best Local Similarity 85.7%; Pred. No. 1.9;
Matches 30; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 16 TTTTATTTTTTTTTTGGGAATATAAAGCGCGC 50
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 31 TTTTATTTTTTTTTTGGCATATAAATAGGCTGC 65

RESULT 4
US-07-750-080A-12/c
; Sequence 12, Application US/07750080A
; Patent No. 5445953
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF A MODIFIED
; TITLE OF INVENTION: EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

Query Match          36.0%; Score 27; DB 1; Length 63;
Best Local Similarity 85.7%; Pred. No. 1.8;
Matches 30; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 16 TTTTATTTTTTTTTTGGGAATATAAAGCGCGC 50
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11 TTTTATTTTTTTTTTGGCATATAAATAGGCTGC 45

RESULT 3
US-07-750-080A-11
; Sequence 11, Application US/07750080A
; Patent No. 5445953
```

QY 16 TTTTATTTTTTTTGGAAATATAAGCGGGCCGCATGGGGCCCGGCCCAACGGC 72

Db 85 TTTTTCATGCTGGCGGGAAGGC 29

RESULT 7
US-07-750-080A-14
; Sequence 14, Application US/07750080A
; Patent No. 5445953
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF A MODIFIED
; TITLE OF INVENTION: EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07750.080A
; FILING DATE: 19910826
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/106 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 134 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: pN29pt-S4
; US-07-750-080A-14

Query Match 35.5%; Score 26.6; DB 1; Length 134;
Best Local Similarity 66.7%; P-red. No. 2.9;
Matches 38; Conservative 0; Mismatches 19; Indels 0; Gaps 0

QY 16 TTTTATTTTTTTTTTGGAATAATAACGCCGCCCATGGCCCGCGCCCAACGCGC 72
|||||
Db 29 TTTTTCATGCTGGCGGGAAGGC 85

RESULT 8
US-08-747-121-20
; Sequence 20, Application US/08747121
; Patent No. 5874290
; GENERAL INFORMATION:
; APPLICANT: Murphy, Gerald
; APPLICANT: Boynton, Alton
; APPLICANT: Sehgal, Anil
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID
; TITLE OF INVENTION: SEQUENCES OF A D2-2 GENE ASSOCIATED WITH
; TITLE OF INVENTION: BRAIN TUMORS AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:

```
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Baldwin, Geraldine F
/ REGISTRATION NUMBER: 31,232
/ REFERENCE/DOCKET NUMBER: 8511-008
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212)7909090
/ TELEFAX: (212)8698864
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2002 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
US-08-747-121-1

Query Match 34.7%; Score 26; DB 3; Length 2002;
Best Local Similarity 70.0%; Pred. No. 10;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 16 TTTTATTTTGTGGAATATAAGCGCGCCGCGCCATGGCCCGCGCCG 65
Db 1021 TTTTGTGTTTTTTTAAAGACAGAGAGTCTGCTGCGCCGCGCGC 1070

RESULT 10
US-08-715-568A-2/c
/ Sequence 2, Application US/08715568A
/ Patent No. 5856463
/ GENERAL INFORMATION:
/ APPLICANT: Prydz, Hans Peter Blankenborg
/ APPLICANT: Prydz, Gaute
/ TITLE OF INVENTION: PSKH-1 Ribozymes and Uses in Disease
/ TITLE OF INVENTION: Treatment
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik
/ STREET: 600 South Avenue West
/ CITY: Westfield
/ STATE: NJ
/ COUNTRY: USA
/ ZIP: 07090-1497
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/715,568A
/ FILING DATE:
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Foley, Shawn P.
/ REFERENCE/DOCKET NUMBER: FORSK 3.0-002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 908-654-5000
/ TELEFAX: 908-654-7866
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3471 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
US-08-715-568A-2

Query Match 33.9%; Score 25.4; DB 3; Length 3471;
Best Local Similarity 68.6%; Pred. No. 18;
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```

```
QY 16 TTTTATTTTGTGGAATATAAGCGCGCCGCGCCATGGCCCGCGCCG 66
Db 3471 TTTTATTTTGTGGAATGAATGGCAGCCCTCCCTGGAGAGCGACCC 3421

RESULT 11
US-07-750-080A-38
/ Sequence 38, Application US/07750080A
/ Patent No. 5445953
/ GENERAL INFORMATION:
/ APPLICANT: DORNER, F.
/ APPLICANT: SCHEIFLINGER, F.
/ APPLICANT: FALKNER, F. G.
/ TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF A MODIFIED
/ TITLE OF INVENTION: EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
/ NUMBER OF SEQUENCES: 42
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: 1800 Diagonal Road, Suite 500
/ CITY: Alexandria
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22313-0299
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/750,080A
/ FILING DATE: 19910826
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 30472/106 IMMU
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703)836-9300
/ TELEFAX: (703)683-4109
/ TELEX: 899149
/ INFORMATION FOR SEQ ID NO: 38:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 55 base pairs
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ IMMEDIATE SOURCE:
/ CLONE: P-artP(11)
US-07-750-080A-38

Query Match 33.6%; Score 25.2; DB 1; Length 55;
Best Local Similarity 90.0%; Pred. No. 6.1;
Matches 27; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 16 TTTTATTTTGTGGAATATAAGCG 45
Db 25 TTTTATTTTGTGGAATATAAGCG 54

RESULT 12
US-07-750-080A-39/c
/ Sequence 39, Application US/07750080A
/ Patent No. 5445953
/ GENERAL INFORMATION:
/ APPLICANT: DORNER, F.
/ APPLICANT: SCHEIFLINGER, F.
/ APPLICANT: FALKNER, F. G.
/ TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF A MODIFIED
/ TITLE OF INVENTION: EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
/ NUMBER OF SEQUENCES: 42
/ CORRESPONDENCE ADDRESS:
```

```

1  PRIOR APPLICATION DATA: including application 1
2  PRIOR APPLICATION DATA: described below:
3  APPLICATION NUMBER: 08/367,395
4  FILING DATE: 12/30/94
5  ATTORNEY/AGENT INFORMATION:
6  NAME: Guise, Jeffrey W.
7  REGISTRATION NUMBER: 34, 613
8  REFERENCE/DOCKET NUMBER: 212/127
9  TELECOMMUNICATION INFORMATION:
10 TELEPHONE: (619) 552-8400
11 TELEFAX: (619) 552-0159
12 TELEX: 67-3510
13 TELEX: SEQUENCE LISTING
14 INFORMATION FOR SEQ ID NO: 7:
15 SEQUENCE CHARACTERISTICS:
16 LENGTH: 3095 base pairs
17 TYPE: nucleic acid
18 STRANDEDNESS: single
19 TOPOLOGY: linear
20 TOPOLOGY: DESCRIPTION:
21 FEATURE:
22 NAME/KEY: Coding Sequence
23 LOCATION: 85....2400
24 US-08-434-000A-7
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
99
```

```

RESULT 14
US-08-968-751-3/c
; Sequence 3, Application US/08958751
; Patent No. 5948643
; GENERAL INFORMATION:
; APPLICANT: Rubinfield, Bonnie
; APPLICANT: Polakis, Paul G.
; APPLICANT: Ligenfelter, Carol
; APPLICANT: Vuong, Terilyn I.
; TITLE OF INVENTION: MODULATORS OF BRCA1 ACTIVITY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ONYX Pharmaceuticals, Inc.
; STREET: 3031 Research Drive
; CITY: Richmond
; STATE: CA
; COUNTRY: USA
; ZIP: 94806
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,751
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Giotta, Gregory
; REGISTRATION NUMBER: 32,028
; REFERENCE/DOCKET NUMBER: ONYX1024 GG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 262-8710
; TELEFAX: (510) 222-9758
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3256 base pairs
; TYPE: nucleic acid

```

```

1  ADDRESSEE:  Foley & Lardner
2  STREET:  1800 Diagonal Road, Suite 500
3  CITY:  Alexandria
4  STATE:  VA
5  COUNTRY:  USA
6  ZIP:  22313-0299
7  COMPUTER READABLE FORM:
8  MEDIUM TYPE:  Floppy disk
9  COMPUTER:  IBM PC compatible
10 OPERATING SYSTEM:  PC-DOS/MS-DOS
11 SOFTWARE:  Patentin Release #1.0, Version #1.25
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER:  US/07/750,080A
14 FILING DATE:  19910826
15 CLASSIFICATION:  435
16 ATTORNEY/AGENT INFORMATION:
17 NAME:  BENT, Stephen A.
18 REGISTRATION NUMBER:  29,768
19 REFERENCE/DOCKET NUMBER:  30472/106 IMMU
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE:  (703)836-9300
22 TELEFAX:  (703)683-4109
23 TELEX:  895149
24 INFORMATION FOR SEQ ID NO: 39:
25 SEQUENCE CHARACTERISTICS:
26 LENGTH:  55 base pairs
27 TYPE:  NUCLEIC ACID
28 STRANDEDNESS:  single
29 TOPOLOGY:  linear
30 MOLECULE TYPE:  DNA (genomic)
31 IMMEDIATE SOURCE:
32 CLONE:  P-artp(12)
33
34 US-07-750-080A-39
35
36 Query Match 33.6%; Score 25.2; DB 1; Length 55;
37 Best Local Similarity 90.0%; Pred. No. 6.1;
38 Matches 27; Conservative 0; Mismatches 3; Indels
39
40 QY 16 TTTTATTTTTTTTTTTGGATATAAAGC 45
41 |||||TTTTTTTTTTTTTTTTTTTTTTTT
42
43 Db 35 TTTTATTTTTTTTTTTGGCATATAAATCG 6
44
45 RESULT 13
46 US-08-434-000A-7/C
47 Sequence 7, Application US/08434000A
48 Patent No. 6046037
49 GENERAL INFORMATION:
50 APPLICANT:  ANDREW C. HIATT, JULIAN
51 APPLICANT:  K.-C. MA, THOMAS LEHNER
52 TITLE OF INVENTION:  IMMUNOGLOBULINS CONTAINING PROTECTION
53 TITLE OF INVENTION:  PROTEINS IN PLANTS AND THEIR USES
54 NUMBER OF SEQUENCES:  19
55 CORRESPONDENCE ADDRESS:
56 ADDRESSEE:  Lyon & Lyon
57 STREET:  633 West Fifth Street
58 CITY:  Suite 4700
59 CITY:  Los Angeles
60 STATE:  California
61 COUNTRY:  U.S.A.
62 ZIP:  90071
63 COMPUTER READABLE FORM:
64 MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
65 MEDIUM TYPE:  storage
66 COMPUTER:  IBM compatible
67 OPERATING SYSTEM:  IBM P.C. DOS 5.0
68 SOFTWARE:  Word Perfect 5.1
69 CURRENT APPLICATION DATA:
70 APPLICATION NUMBER:  US/08/434,000A
71 FILING DATE:
72 CLASSIFICATION:  435
73 PRIOR APPLICATION DATA:

```


STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 34...2541
US-08-968-751-3

Search completed: May 29, 2000, 22:08:54
Job time: 38735 sec

Query Match 33.3%; Score 25; DB 4; Length 3256;
Best Local Similarity 84.8%; Pred. No. 24;
Matches 28; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 10 TTGAATTTTATTTTGTGAATATAA 42
DB 3237 TCGAGTTTTTGTGATATAA 3205

RESULT 15
US-08-306-691B-16/c
Sequence 16, Application US/08306691B
Patent No. 5734039
GENERAL INFORMATION:
APPLICANT: Calabretta, Bruno
APPLICANT: Skorski, Tomasz
TITLE OF INVENTION: ANTISENSE
TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seidel, Gonda, Lavoigne & Monaco, P.C.
STREET: Two Penn Center, Suite 1800
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19102

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306.691B
FILING DATE: September 15, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5734039e
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2436 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-306-691B-16

Query Match 33.1%; Score 24.8; DB 2; Length 2436;
Best Local Similarity 67.3%; Pred. No. 25;
Matches 35; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GGCCAAAATGAAATTTTATTTTGTGAATATAAAGCGGCCCA 52
DB 1894 GACTCAAAATTCAGTATTCATTTTGTGATATAAAGCGGCCCA 1843

us-08-935-377-3.rni

Wed May 31 10:04:52 2000

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2000, 09:48:19 ; Search time 2276.24 Seconds
(without alignments)
147.011 Million cell updates/sec

Title: US-08-935-377-3
Perfect score: 75
Sequence: 1 GCGCAAAATGGAATTGA.....GCGCGCCGCAACGCGGA 75

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5142629 seqs, 2230885800 residues
Total number of hits satisfying chosen parameters: 10285240

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA:*

1: /cgnl_6/ptodata/1/pna/PCTUS_COMB.seq:*

2: /cgnl_6/ptodata/1/pna/US06_COMB.seq:*

3: /cgnl_6/ptodata/1/pna/US07_COMB.seq:*

4: /cgnl_6/ptodata/1/pna/US08_COMB.seq:*

5: /cgnl_6/ptodata/1/pna/US081A_COMB.seq:*

6: /cgnl_6/ptodata/1/pna/US081B_COMB.seq:*

7: /cgnl_6/ptodata/1/pna/US081C_COMB.seq:*

8: /cgnl_6/ptodata/1/pna/US081D_COMB.seq:*

9: /cgnl_6/ptodata/1/pna/US082A_COMB.seq:*

10: /cgnl_6/ptodata/1/pna/US082B_COMB.seq:*

11: /cgnl_6/ptodata/1/pna/US082C_COMB.seq:*

12: /cgnl_6/ptodata/1/pna/US083A_COMB.seq:*

13: /cgnl_6/ptodata/1/pna/US083B_COMB.seq:*

14: /cgnl_6/ptodata/1/pna/US084A_COMB.seq:*

15: /cgnl_6/ptodata/1/pna/US084B_COMB.seq:*

16: /cgnl_6/ptodata/1/pna/US084C_COMB.seq:*

17: /cgnl_6/ptodata/1/pna/US085_COMB.seq:*

18: /cgnl_6/ptodata/1/pna/US086_COMB.seq:*

19: /cgnl_6/ptodata/1/pna/US087A_COMB.seq:*

20: /cgnl_6/ptodata/1/pna/US087B_COMB.seq:*

21: /cgnl_6/ptodata/1/pna/US087C_COMB.seq:*

22: /cgnl_6/ptodata/1/pna/US088A_COMB.seq:*

23: /cgnl_6/ptodata/1/pna/US088B_COMB.seq:*

24: /cgnl_6/ptodata/1/pna/US088C_COMB.seq:*

25: /cgnl_6/ptodata/1/pna/US088D_COMB.seq:*

26: /cgnl_6/ptodata/1/pna/US089A_COMB.seq:*

27: /cgnl_6/ptodata/1/pna/US089B_COMB.seq:*

28: /cgnl_6/ptodata/1/pna/US089C_COMB.seq:*

29: /cgnl_6/ptodata/1/pna/US089D_COMB.seq:*

30: /cgnl_6/ptodata/1/pna/US089E_COMB.seq:*

31: /cgnl_6/ptodata/1/pna/US089F_COMB.seq:*

32: /cgnl_6/ptodata/1/pna/US089G_COMB.seq:*

33: /cgnl_6/ptodata/1/pna/US089H_COMB.seq:*

34: /cgnl_6/ptodata/1/pna/US089I_COMB.seq:*

35: /cgnl_6/ptodata/1/pna/US090A_COMB.seq:*

36: /cgnl_6/ptodata/1/pna/US090B_COMB.seq:*

37: /cgnl_6/ptodata/1/pna/US090C_COMB.seq:*

38: /cgnl_6/ptodata/1/pna/US090D_COMB.seq:*

39: /cgnl_6/ptodata/1/pna/US090E_COMB.seq:*

40: /cgnl_6/ptodata/1/pna/US091A_COMB.seq:*

41: /cgnl_6/ptodata/1/pna/US091B_COMB.seq:*

42: /cgnl_6/ptodata/1/pna/US091C_COMB.seq:*

43: /cgnl_6/ptodata/1/pna/US092A_COMB.seq:*

44: /cgnl_6/ptodata/1/pna/US092B_COMB.seq:*

45: /cgnl_6/ptodata/1/pna/US092C_COMB.seq:*

46: /cgnl_6/ptodata/1/pna/US092D_COMB.seq:*

47: /cgnl_6/ptodata/1/pna/US092E_COMB.seq:*

48: /cgnl_6/ptodata/1/pna/US093A_COMB.seq:*

49: /cgnl_6/ptodata/1/pna/US093B_COMB.seq:*

50: /cgnl_6/ptodata/1/pna/US093C_COMB.seq:*

51: /cgnl_6/ptodata/1/pna/US093D_COMB.seq:*

52: /cgnl_6/ptodata/1/pna/US093E_COMB.seq:*

53: /cgnl_6/ptodata/1/pna/US093F_COMB.seq:*

54: /cgnl_6/ptodata/1/pna/US094A_COMB.seq:*

55: /cgnl_6/ptodata/1/pna/US094B_COMB.seq:*

56: /cgnl_6/ptodata/1/pna/US094C_COMB.seq:*

57: /cgnl_6/ptodata/1/pna/US095_COMB.seq:*

58: /cgnl_6/ptodata/1/pna/US6000_COMB.seq:*

59: /cgnl_6/ptodata/1/pna/US6001A_COMB.seq:*

60: /cgnl_6/ptodata/1/pna/US6001B_COMB.seq:*

61: /cgnl_6/ptodata/1/pna/US6002A_COMB.seq:*

62: /cgnl_6/ptodata/1/pna/US6002B_COMB.seq:*

63: /cgnl_6/ptodata/1/pna/US6002C_COMB.seq:*

64: /cgnl_6/ptodata/1/pna/US6003A_COMB.seq:*

65: /cgnl_6/ptodata/1/pna/US6003B_COMB.seq:*

66: /cgnl_6/ptodata/1/pna/US6004A_COMB.seq:*

67: /cgnl_6/ptodata/1/pna/US6004B_COMB.seq:*

68: /cgnl_6/ptodata/1/pna/US6004C_COMB.seq:*

69: /cgnl_6/ptodata/1/pna/US6005_COMB.seq:*

70: /cgnl_6/ptodata/1/pna/US6006_COMB.seq:*

71: /cgnl_6/ptodata/1/pna/US6007_COMB.seq:*

72: /cgnl_6/ptodata/1/pna/US6008A_COMB.seq:*

73: /cgnl_6/ptodata/1/pna/US6008B_COMB.seq:*

74: /cgnl_6/ptodata/1/pna/US6008C_COMB.seq:*

75: /cgnl_6/ptodata/1/pna/US6009A_COMB.seq:*

76: /cgnl_6/ptodata/1/pna/US6009B_COMB.seq:*

77: /cgnl_6/ptodata/1/pna/US6010A_COMB.seq:*

78: /cgnl_6/ptodata/1/pna/US6010B_COMB.seq:*

79: /cgnl_6/ptodata/1/pna/US6011A_COMB.seq:*

80: /cgnl_6/ptodata/1/pna/US6011B_COMB.seq:*

81: /cgnl_6/ptodata/1/pna/US6012A_COMB.seq:*

82: /cgnl_6/ptodata/1/pna/US6012B_COMB.seq:*

83: /cgnl_6/ptodata/1/pna/US6013A_COMB.seq:*

84: /cgnl_6/ptodata/1/pna/US6013B_COMB.seq:*

85: /cgnl_6/ptodata/1/pna/US6014A_COMB.seq:*

86: /cgnl_6/ptodata/1/pna/US6014B_COMB.seq:*

87: /cgnl_6/ptodata/1/pna/US6014C_COMB.seq:*

88: /cgnl_6/ptodata/1/pna/US6015_COMB.seq:*

89: /cgnl_6/ptodata/1/pna/US6016A_COMB.seq:*

90: /cgnl_6/ptodata/1/pna/US6016B_COMB.seq:*

91: /cgnl_6/ptodata/1/pna/US6016C_COMB.seq:*

92: /cgnl_6/ptodata/1/pna/US6017A_COMB.seq:*

93: /cgnl_6/ptodata/1/pna/US6017B_COMB.seq:*

94: /cgnl_6/ptodata/1/pna/US6017C_COMB.seq:*

95: /cgnl_6/ptodata/1/pna/US6018A_COMB.seq:*

96: /cgnl_6/ptodata/1/pna/US6018B_COMB.seq:*

97: /cgnl_6/ptodata/1/pna/US6018C_COMB.seq:*

98: /cgnl_6/ptodata/1/pna/US6019_COMB.seq:*

99: /cgnl_6/ptodata/1/pna/PCT_NEW_COMB.seq:*

100: /cgnl_6/ptodata/1/pna/US06_NEW_COMB.seq:*

101: /cgnl_6/ptodata/1/pna/US07_NEW_COMB.seq:*

102: /cgnl_6/ptodata/1/pna/US08_NEW_COMB.seq:*

103: /cgnl_6/ptodata/1/pna/US09_NEW_COMB.seq:*

104: /cgnl_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	75	100.0	75	28	US-08-935-377-3	Sequence 3, Appl
2	59	78.7	59	28	US-08-935-377-25	Sequence 25, Appl

us-08-935-377-3.rnp

wed May 31 10:04:52 2000

NAME: Steffe, Eric K
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 75 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
NAME/KEY: CDS
LOCATION: 52..75
US-08-935-377-3

Query Match 100.0%; Score 75; DB 28; Length 75;
Best Local Similarity 100.0%; Pred. No. 5.3e-09; Indels 0; Gaps 0;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAAAATGAATTTATTTTGGATATAAAGCGCGCCATGGCCCG 60
Db 1 GCCCAAAATGAATTTATTTTGGATATAAAGCGCGCCATGGCCCG 60
QY 61 GCCGCAACGGCGGA 75
Db 61 GCCGCAACGGCGGA 75

RESULT 2
US-08-935-377-25
SEQUENCE 25, Application US/08935377
GENERAL INFORMATION:
APPLICANT: Zauderer, Maurice
TITLE OF INVENTION: T Cells Specific for Target Antigens and
TITLE OF INVENTION: Vaccines Based Thereon
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/935,377
FILING DATE: 22-SEP-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 59 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-935-377-25

Sequence 26, Appl
Sequence 32, Appl
Sequence 8, Appl
Sequence 9, Appl
Sequence 10, Appl
Sequence 66, Appl
Sequence 66, Appl
Sequence 66, Appl
Sequence 72, Appl
Sequence 72, Appl
Sequence 72, Appl
Sequence 70, Appl
Sequence 70, Appl
Sequence 33, Appl
Sequence 67, Appl
Sequence 67, Appl
Sequence 11, Appl
Sequence 12, Appl
Sequence 56, Appl
Sequence 56, Appl
Sequence 56, Appl
Sequence 57, Appl
Sequence 57, Appl
Sequence 57, Appl
Sequence 65, Appl
Sequence 65, Appl
Sequence 1, Appl
Sequence 416, Ap
Sequence 4576, Ap
Sequence 372, App
Sequence 32, Appl
Sequence 1517, Ap
Sequence 1517, Ap
Sequence 4227, Ap
Sequence 33060, A
Sequence 2413, Ap
Sequence 9281, Ap
Sequence 1777, Ap
Sequence 4913, Ap
Sequence 18207, A

US-08-935-377-26
US-08-935-377-32
US-08-322-358-8
US-08-322-358-9
US-08-322-358-10
US-07-914-7388-66
US-08-358-928-66
US-08-651-472-66
US-07-914-7388-72
US-08-358-928-72
US-08-651-472-72
US-07-914-7388-70
US-08-358-928-70
US-08-651-472-70
US-08-935-377-33
US-07-914-7388-67
US-08-358-928-67
US-08-651-472-67
US-08-322-358-11
US-08-322-358-12
US-08-322-358-12
US-08-358-928-56
US-08-651-472-56
US-07-914-7388-57
US-08-358-928-57
US-08-651-472-57
US-08-651-472-57
US-07-914-7388-65
US-08-358-928-65
US-08-651-472-65
US-08-935-377-1
US-09-288-687-4416
US-09-288-687-4576
US-60-182-569-372
US-60-195-135-32
US-08-196-362A-1517
US-08-196-362D-1517
US-09-401-645-4227
US-09-293-972-33060
US-09-515-126-2413
US-09-293-972-9281
US-09-353-690-1777
US-60-144-351-4913
US-09-321-214-18207

ALIGNMENTS

RESULT 1
US-08-935-377-3
SEQUENCE 3, Application US/08935377
GENERAL INFORMATION:
APPLICANT: Zauderer, Maurice
TITLE OF INVENTION: T Cells Specific for Target Antigens and
TITLE OF INVENTION: Vaccines Based Thereon
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/935,377
FILING DATE: 22-SEP-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:

```
Query Match      78.7%; Score 59; DB 28; Length 59;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAAAATTCGAAATTTATTTTGGGAATATAAGCGCGCCATGGGCC 59
   |||||||
Db 1 GCCCAAAATTCGAAATTTATTTTGGGAATATAAGCGCGCCATGGGCC 59
   |||||||

RESULT 3
US-08-935-377-26/c
; Sequence 26, Application US/08935377
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; TITLE OF INVENTION: T Cells Specific for Target Antigens and
; TITLE OF INVENTION: Vaccines Based Thereon
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,377
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 59 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-935-377-26

Query Match      78.7%; Score 59; DB 28; Length 59;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AAAAATGAAATTTATTTTGGGAATATAAGCGCGCCATGGGCCGGCC 63
   |||||||
Db 59 AAAAATGAAATTTATTTTGGGAATATAAGCGCGCCATGGGCCGGCC 1
   |||||||

RESULT 4
US-08-935-377-32
; Sequence 32, Application US/08935377
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; TITLE OF INVENTION: T Cells Specific for Target Antigens and
; TITLE OF INVENTION: Vaccines Based Thereon
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
```

```
ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,377
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-935-377-32

Query Match      56.5%; Score 42.4; DB 28; Length 77;
Best Local Similarity 97.7%; Pred. No. 0.31;
Matches 43; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAAAATTCGAAATTTATTTTGGGAATATAAGCGCGCCATGGGCC 44
   |||||||
Db 1 GCCCAAAATTCGAAATTTATTTTGGGAATATAAGCGCGCCATGGGCC 44
   |||||||

RESULT 5
US-08-322-358-8
; Sequence 8, Application US/08322358
; GENERAL INFORMATION:
; APPLICANT: FALKNER, Falko-Gunter
; APPLICANT: BODENER, Walter
; APPLICANT: DORNER, Friedrich
; APPLICANT: MOSS, Bernard
; TITLE OF INVENTION: PRODUCTION OF ISOLATED PROTEINACOUS
; TITLE OF INVENTION: MATERIALS USING RECOMBINANT AVIPOX VIRUS VECTORS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/322,358
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/882,768
; FILING DATE:
; APPLICATION NUMBER: US 07/734,741
; FILING DATE: 23-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30357/108 KIST
; TELECOMMUNICATION INFORMATION:
```

US-08-322-358-9

TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: Oligonucleotide oselP2
US-08-322-358-8

Query Match 53.3%; Score 40; DB 12; Length 84;
Best Local Similarity 82.1%; Pred. No. 1.2;
Matches 46; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CCAAAATTGAATTTTATTTTTTTGGAATAAAGCGGCCCATGGCC 58
| | | | |
DB 9 CTAAAAATTGAATTTTATTTTTTTGGAATAAAGCGCCTCCATGGCC 64
| | | | |

RESULT 6
US-08-322-358-9/c
Sequence 9, Application US/08322358
GENERAL INFORMATION:
APPLICANT: FALKNER, Falko-Gunter
APPLICANT: BODMER, Walter
APPLICANT: DORNER, Friedrich
APPLICANT: MOSS, Bernard
TITLE OF INVENTION: PRODUCTION OF ISOLATED PROTEINACEOUS
TITLE OF INVENTION: MATERIALS USING RECOMBINANT AVIOPX VIRUS VECTORS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322.358
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/882.768
FILING DATE:
APPLICATION NUMBER: US 07/734,741
FILING DATE: 23-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30357/108 KIST
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: Plasmid pTKgt-selp (4723bp)
US-08-322-358-10

Query Match 53.3%; Score 40; DB 12; Length 84;
Best Local Similarity 82.1%; Pred. No. 2.3;
Matches 46; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CCAAAATTGAATTTTATTTTTTTGGAATAAAGCGGCCCATGGCC 58
| | | | |
DB 1524 CTAAAAATTGAATTTTATTTTTTTGGAATAAAGCGCCTCCATGGCC 1579
| | | | |

US-08-322-358-9

TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: Oligonucleotide oselP2
US-08-322-358-8

Query Match 53.3%; Score 40; DB 12; Length 84;
Best Local Similarity 82.1%; Pred. No. 1.2;
Matches 46; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CCAAAATTGAATTTTATTTTTTTGGAATAAAGCGGCCCATGGCC 58
| | | | |
DB 9 CTAAAAATTGAATTTTATTTTTTTGGAATAAAGCGCCTCCATGGCC 64
| | | | |

RESULT 6
US-08-322-358-9/c
Sequence 9, Application US/08322358
GENERAL INFORMATION:
APPLICANT: FALKNER, Falko-Gunter
APPLICANT: BODMER, Walter
APPLICANT: DORNER, Friedrich
APPLICANT: MOSS, Bernard
TITLE OF INVENTION: PRODUCTION OF ISOLATED PROTEINACEOUS
TITLE OF INVENTION: MATERIALS USING RECOMBINANT AVIOPX VIRUS VECTORS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322.358
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/882.768
FILING DATE:
APPLICATION NUMBER: US 07/734,741
FILING DATE: 23-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30357/108 KIST
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 86 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: Oligonucleotide oselP2
US-08-322-358-10

Query Match 53.3%; Score 40; DB 12; Length 4723;
Best Local Similarity 82.1%; Pred. No. 2.3;
Matches 46; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CCAAAATTGAATTTTATTTTTTTGGAATAAAGCGGCCCATGGCC 58
| | | | |
DB 1524 CTAAAAATTGAATTTTATTTTTTTGGAATAAAGCGCCTCCATGGCC 1579
| | | | |

us-08-935-377-3.rnp

Wed May 31 10:04:52 2000

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/651,472
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/914,738
; FILING DATE: 20-JUL-1992
; PRIOR APPLICATION DATA: US 07/750,080
; APPLICATION NUMBER: US 07/914,738
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/166/IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other nucleic acid;
; DESCRIPTION: Synthetic DNA oligonucleotide
; IMMEDIATE SOURCE:
; CLONE: pslp-gp160MN
;
; US-08-651-472-66
;
; Query Match 52.8%; Score 39.6; DB 18; Length 6474;
; Best Local Similarity 83.3%; Pred. No. 3; Indels 0; Gaps 0;
; Matches 45; Conservative 0; Mismatches 9;
;
; QY 5 AAAAAATTGAAATTTATTTTGGGAATATAAAGCGCGCCATGGGCC 58
; DB 3964 AAAAAATTGAAATTTATTTTGGGAATATAAAGCGCGCCATGGGCC 3911
;
; RESULT 11
; US-07-914-738B-72/c
; Sequence 72, Application US/07914738B
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; APPLICANT: PELEIDERER, M.
; TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF A MODIFIED
; TITLE OF INVENTION: EUKARYOTIC CYTOPLASMIC DNA VIRUS GENOME
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/914,738B
; FILING DATE: 19920720
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/750,080
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/166/IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5532 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other nucleic acid;
; DESCRIPTION: Synthetic DNA oligonucleotide
; IMMEDIATE SOURCE:
; CLONE: pN2gpta-FIX
;
; US-07-914-738B-72
;
; Query Match 51.7%; Score 38.8; DB 3; Length 5532;
; Best Local Similarity 79.3%; Pred. No. 4.6;
; Matches 46; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
;
; QY 3 CCAAAATTGAAATTTATTTTGGGAATATAAAGCGCGCCATGGGCCG 60
; DB 3709 CTAATAATTGAAATTTATTTTGGGAATATAAAGCGCGCCATGGGCCG 3652
;
; RESULT 12
; US-08-358-928-72/c
; Sequence 72, Application US/08358928
; GENERAL INFORMATION:
; APPLICANT: DORNER, Friedrich
; APPLICANT: SCHEIFLINGER, Friedrich
; APPLICANT: FALKNER, Falko Gunter
; APPLICANT: PELEIDERER, Michael
; TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF CHIMERIC
; TITLE OF INVENTION: VIRUSES CONTAINING HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/358,928
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/914,738
; FILING DATE: 20-JUL-1992
; PRIOR APPLICATION DATA: US 07/750,080
; APPLICATION NUMBER:
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/166/IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:

```


US-08-651-472-72

us-08-935-377-3.rnp

Wed May 31 10:04:52 2000

```

; Sequence 70, Application US/08358928
; GENERAL INFORMATION:
; APPLICANT: DORNER, Friedrich
; APPLICANT: SCHEIFLINGER, Friedrich
; APPLICANT: FALKNER, Falko Gunter
; APPLICANT: PFLEIDERER, Michael
; TITLE OF INVENTION: DIRECT MOLECULAR CLONING OF CHIMERIC
; TITLE OF INVENTION: VIRUSES CONTAINING HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
; TITLE OF INVENTION: (HIV-1) ANTIGENS
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/358,928
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/914,738
; FILING DATE: 20-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/750,080
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/166/IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other nucleic acid;
; DESCRIPTION: Synthetic DNA oligonucleotide
; IMMEDIATE SOURCE:
; CLONE: self promoter
;
US-08-358-928-70

```

Query Match 51.2%; Score 38.4; DB 13; Length 49;
 Best Local Similarity 97.5%; Pred. No. 2.6;
 Matches 39; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 3 CCAAAATTTGAATTTATTTTATTTTGGGAATATAAA 42
   | | | | | | | | | | | | | | | | | | | |
Db 9 CTAAATTTGAATTTATTTTATTTTGGGAATATAAA 48

```

Search completed: May 30, 2000, 09:48:27
 Job time: 60177 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:13:10 ; Search time 2192.43 Seconds
(without alignments)
138.655 Million cell updates/sec

Title: US-08-935-377-3
Perfect score: 75
Sequence: 1 GCCCAAAATGAAATTTTA.....GCCCGCGCCCAACGCCGGA 75

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_est1:*
- 2: em_est2:*
- 3: em_est3:*
- 4: em_est4:*
- 5: em_est5:*
- 6: em_est6:*
- 7: em_est7:*
- 8: em_est8:*
- 9: em_est9:*
- 10: em_est10:*
- 11: em_est11:*
- 12: em_est12:*
- 13: em_est13:*
- 14: em_est14:*
- 15: em_est15:*
- 16: em_est16:*
- 17: em_est17:*
- 18: em_est18:*
- 19: em_est19:*
- 20: gb_est1:*
- 21: gb_est2:*
- 22: gb_est3:*
- 23: gb_est4:*
- 24: gb_est5:*
- 25: gb_est6:*
- 26: gb_est7:*
- 27: gb_est8:*
- 28: gb_est9:*
- 29: gb_est10:*
- 30: gb_est11:*
- 31: gb_est12:*
- 32: gb_est13:*
- 33: gb_est14:*
- 34: gb_est15:*
- 35: gb_est16:*
- 36: gb_est17:*
- 37: gb_est18:*
- 38: gb_est19:*
- 39: gb_est20:*
- 40: gb_est21:*
- 41: gb_est22:*
- 42: gb_est23:*
- 43: gb_est24:*

- 44: gb_est25:*
- 45: gb_est26:*
- 46: gb_est27:*
- 47: gb_est28:*
- 48: gb_est29:*
- 49: gb_est30:*
- 50: em_est21:*
- 51: em_est22:*
- 52: gb_est30:*
- 53: gb_est31:*
- 54: gb_est32:*
- 55: em_est23:*
- 56: em_est24:*
- 57: em_est25:*
- 58: em_est26:*
- 59: gb_est33:*
- 60: gb_est34:*
- 61: gb_est35:*
- 62: gb_est36:*
- 63: gb_est37:*
- 64: gb_est38:*
- 65: em_est27:*
- 66: em_est28:*
- 67: em_est29:*
- 68: em_est30:*
- 69: gb_est39:*
- 70: gb_est40:*
- 71: gb_est41:*
- 72: gb_est42:*
- 73: gb_est43:*
- 74: gb_est44:*
- 75: em_est31:*
- 76: em_est32:*
- 77: em_est33:*
- 78: em_est34:*
- 79: gb_est45:*
- 80: gb_est46:*
- 81: gb_est47:*
- 82: gb_gss1:*
- 83: gb_gss2:*
- 84: gb_gss3:*
- 85: gb_gss4:*
- 86: em_gss1:*
- 87: em_gss2:*
- 88: em_gss3:*
- 89: em_gss4:*
- 90: gb_gss5:*
- 91: gb_gss6:*
- 92: em_gss5:*
- 93: em_gss6:*
- 94: gb_gss7:*
- 95: gb_gss8:*
- 96: gb_gss9:*
- 97: em_gss7:*
- 98: em_gss8:*
- 99: gb_gss11:*
- 100: gb_gss10:*
- 101: em_gss9:*
- 102: em_gss10:*
- 103: em_gss11:*
- 104: em_gss12:*
- 105: gb_gss12:*
- 106: gb_gss13:*
- 107: gb_gss14:*
- 108: gb_gss15:*
- 109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> -. This Drosophila melanogaster BAC library (DrosBAC) was made by Alain Billaud at CDPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelosAC11.

```

FEATURES
Location/Qualifiers
1..1201
/organism="Drosophila melanogaster"
/plasmid="pBelorAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN15M07"
/note="end : T7"

```

BASE COUNT	286 a	290 c	192 g	283 t	150 others
ORIGIN					
Query Match			47.2%;	Score 35.4;	DB 83; Length 1201;
Best Local similarity			47.5%;	Pred. No. 7.8;	

[illegible][illegible][illegible]

LOCUS	AW297017.1	138 bp
DEFINITION	UT-H-BWO-af-f-0-0-UI.s1 NCI_CGAP_Sub6 Homo sapiens cDNA clone IMAGE:2731457 3', mRNA sequence.	
ACCESSION	AW297017	
VERSION	AW297017.1	GI:6703653

KEYWORDS EST. human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Cararrhini; Hominiidae; Homo.

REFERENCE
1 (bases 1 to 158)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE
Tumor Gene Index

Unpublished (1997)
On Dec 20, 1995 this sequence version replaced gi:1133426.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert_Straussberg@mail.nih.gov

NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: www.bio.llnl.gov/obrp/image.html The following repetitive elements were found in this cDNA sequence: 99-142,

```
>GC rich#Low_complexity
Seg primer: M13 Forward
POLYA-Yes.
FEATURES             Location/Qualifiers
     source            1..158
```

```

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:2731457"
/clone_lib="NCI_CGAP_Sub6"

```

```

/lab_host="Dhl0B (Life Technologies,
/notes=vector: pT7r3D-Pac (Pharmacia) with a modified

```

Result	No.	Score	Query Match	Length	DB	ID	Description
c	1	35.4	47.2	1201	83	CNS01651	Al106303 Drosophil
	2	33.4	44.5	158	79	AW297017	AW297017 UI-H-BW0-
	3	30.6	40.8	1101	82	CNS00Y65	Al096911 Drosophil
c	3	30.4	40.5	909	85	AQ743767	AQ743767 HS_5501_A
	4	30.4	40.5	909	85	AQ743767	AQ743767 HS_5501_A
	4	30.4	40.5	909	85	AQ743767	AQ743767 HS_5501_A

[illegible]

C 13	28.8	38.4	378	28	AA123454	mp3aao02.1
C 14	28.8	38.4	423	53	A1683617	7529055.x
C 15	28.8	38.4	425	26	W30577	mc28a08.r1
C 16	28.8	38.4	464	69	AW125685	UI-M-BH2.
C 17	28.8	38.4	553	70	AW155524	mg1e0030P
C 18	28.8	38.4	805	90	A0782146	HS_3174_B
				20	A0702215	Drosophila
				20	A1058326	Drosophila

C	19	28.8	38.4	1062	82	CNS000AA	AU073906	AU073906	AU073906
C	20	28.6	38.1	194	54	AU073906	qu81d09.x	qu81d09.x	qu81d09.x
C	21	28.6	38.1	372	44	A1285735	Drosophila	Drosophila	Drosophila
C	22	28.6	38.1	953	82	CNS002MH	AL097859	AL097859	AL097859
C	23	28.6	38.1	1201	83	CNS0162C	AL106206	AL106206	AL106206

[illegible]

	28	28.4	37.9	1101	63	CN178361	AU783608
	c	29	28.2	37.6	184	60	AI783861
	c	30	28.2	37.6	409	54	AI788954
	c	31	28.2	37.6	498	34	AA536039
	c	32	28.2	37.6	720	36	AA560391
							00256 MUR
							nj81c08. s
							tj81c08. x
							tr29c08. x
							tr29c08. x

[illegible]

37	B20040 T7B22-T7.2
38	A607032
39	A607031
40	B20040 T7B22-T7.2
41	AL074548 Drosophil
42	AW074018 xb06d08.x
43	AW074018 xb06d08.x
44	AW445994 tJ32c09.x

41	27.8	37.1	389	38	AA773220	af78604.r
42	27.8	37.1	522	52	AA773220	486041F08
43	27.8	37.1	532	52	AI629893	AI629893
44	27.8	37.1	531	40	AA947860	AA947860
45	27.8	37.1	602	105	AQ320590	AQ320590

ALIGNMENTS

RESULT	1					
CNS01651/c						
LOCUS	CNS01651	1201 bp	DNA	GSS	26-JUL-1999	
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC					

ACCESSION AL106303
VERSION AL106303.1 GI:5621247

KEYWORDS
SOURCE
ORGANISM
fruit fly.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
pterivota; Neoptera; Endopterygota; Diptera; Brachycera;

REFERENCE
1 (bases 1 to 1201)
AUTHORS
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
Genoscope.
Direct Submission
TITLE -
Centre National de Sequencage

JOURNAL
Submitted (23-JUL-1999) genoscope
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)

Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert.Strausberg@nih.gov

The sequence contained an oligo-dn track that was present in the polynucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information: M.B. Soares Lab Clone distribution: I.M.A.G.E. Consortium/LNLN at:
www.bio.lnl.gov/brrp/Image.html

The following repetitive elements were found in this cdna sequence: 15-66, >MLT1E#LTR/MALR

Seq primer: M13 Forward

PolyA=yes.

Location/Qualifiers

1. .468

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2732001"

/Clone_lib="NCI_CGAP_Sub6"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pRTT3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not i; Site.2: Eco RI; NCI_CGAP_Sub6 is a subtracted library derived from BW, which consists of a mixture of four normalized libraries: NCI_CGAP_Brn50, NCI_CGAP_Lu13, NCI_CGAP_Ov18, GBC1. The NCI_CGAP_Sub6 library had 7 million recombinants. A single-stranded DNA preparation of BW was used as a tracer in a subtractive hybridization with a driver comprising: the IMAGE pool (NCI_CGAP_Kid3 pool 1 LLAM 3334-3337; 3682-3683, 3798-3803 (IMAGE Clonoids 132376-132391), 1456008-1456775, 1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clonoids 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE Clonoids 1414820-1417991, 1520804-1522439); NCI_CGAP_GC4 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids 985608-986759, 1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE Clonoids 1057416-1061255, 1144584-1145351). (50% of the driver population), plus a pool of 3,840 arrayed clones from NCI_CGAP_Sub1 (IMAGE Clonoids 2708616-2710535) and NCI_CGAP_Sub2 (IMAGE Clonoids 2710536-2712455) (20% of the driver population), plus a pool of 11,136 clones from NCI_CGAP_Sub3 (IMAGE Clonoids 2712456-2723591) (30% of the driver population). Subtraction was performed as previously described (Bonaldi, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806. TAG_LIB=NCI_CGAP-Lu13 TAG_TISSUE=lung TAG_SEQ=GCGG"

109 a lll c 88 g 160 t

39.2%; Score 29.4; DB 79; Length 468;
Similarity 63.4%; Pred. No. 3.2e+02;
45; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

CCCAAAATTTGAATTTATTATTTTTTTTGGAATATAAGCGGCCCATGGCCCCG 60
||||| |||| | ||| | ||| | ||| | ||| | ||| | ||| |
TCACGATTGACTTTTCAGTTGTGGCTGGGGACTTTAAGAGGCGCCTCCTAGGCCT 103
CGCCACGG 71
||| ||
CTCCTGTGG 92

Query Match 39.2%; Score 29.4; DB 79; Length 468;
Best Local Similarity 63.4%; Pred. No. 3.2e+02;
Matches 45; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 GGCCAAAAATTGAAATTTTATTTTTTTTTTTGGGAATATAACGGCGGCCCATGGGCCG 60
||| ||||| ||| | ||| | ||| | ||| | ||| |
Db 162 GGTCCAGGATTCACCTTTTCAGGTTTGCGTGGGGACTTAAGAGGGCTTCATGASCCCT 103

QY 61 GCCGCCAACGG 71

Db 102 GCCTCCTGTGG 92

RESULT 9
AA339067/c

Wed May 31 10:04:53 2000

```

LOCUS      AA339067      230 bp      mRNA      EST      21-APR-1997
DEFINITION EST44123 Fetal brain 1 Homo sapiens cDNA 5' end similar to EST
            containing Alu repeat, mRNA sequence.
ACCESSION  AA339067
VERSION     AA339067.1 GI:1991315
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 230)
AUTHORS     Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodex,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkley,S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
            Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
            Nature 377 (6547 Suppl), 3-174 (1995)
JOURNAL     12140200
MEDLINE     On Sep 12, 1996 this sequence version replaced gi:1393066.
COMMENT     Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
            Location/Qualifiers
                1..230
                /organism="Homo sapiens"
                /db_xref="ATCC (inhost):140684"
                /db_xref="taxon:9606"
                /clone_lib="Fetal brain 1"
                /sex="female"
                /dev_stage="fetus, 24 wks"
                /note="Organ: brain; Vector: pBluescript SK-; Site_1:
                ECORI; Site_2: XhoI"
BASE COUNT  52 a 67 c 48 g 62 t 1 others
ORIGIN
Query Match 38.9%; Score 29.2; DB 32; Length 230;
Best Local Similarity 81.0%; Pred. No. 4.7e+02;
Matches 34; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GGCCAAATGAATTTATTTTATTTTGGATATATAAGCGGCCCATGGCCCG 42
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 185 GGCCAAATGAATTTATTTTATTTTGGATATATAAGCGGCCCATGGCCCG 144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
LOCUS      AI902252/c      644 bp      mRNA      EST      01-DEC-1999
DEFINITION QV-BT003-041198-029 BT003 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AI902252
VERSION     AI902252.1 GI:6492730
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 644)
AUTHORS     HCGP http://www.ludwig.org.br/ORESTES.
            The FAPESP/LICR Human Cancer Genome Project
            Unpublished (1999)
            On Dec 20, 1995 this sequence version replaced gi:1124918.
            Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?l1-QV&t2-QV-BT003-029.html&t3-041198&t4-1)
            Seq primer: puc 18 forward.
            Location/Qualifiers
                1..644
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="BT003"
                /sex="female"
                /dev_stage="Adult"
                /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                SmaI; A mini-library was made by cloning products derived
                from ORESTES PCR (U.S. Letters Patent application No.
                196,716 - Ludwig Institute for Cancer Research) profiles
                into the puc 18 vector. Reverse transcription of tissue
                mRNA and cDNA amplification were performed under low
                stringency conditions. 256 t 19 others
BASE COUNT  141 a 112 c 116 g 256 t 19 others
ORIGIN
Query Match 38.9%; Score 29.2; DB 62; Length 644;
Best Local Similarity 62.2%; Pred. No. 3.2e+02;
Matches 46; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1 GGCCAAATGAATTTATTTTATTTTGGATATATAAGCGGCCCATGGCCCG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 270 GGCCAAATGAATTTATTTTATTTTGGATATATAAGCGGCCCATGGCCCG 211
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 GCGGCCCAACGGCGG 74
    | | | | |
Db 210 GGGGGGGGGGGGGG 197
    | | | | |

RESULT 11
LOCUS      CNS00F3L      1024 bp      DNA      GSS      04-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC:
            BACR30J05 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL069994
VERSION     AL069994.1 GI:4950125
KEYWORDS    GSS.
SOURCE      fruit fly.
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
            Mus 1 to 1024)
            Genoscope.
            Direct Submission
            Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

```


FEATURES	source	Location/Qualifiers
1..368		
		organism="Homo sapiens"
		db_xref="taxon:9606"
		/clone="IMAGE:2532498"
		/clone_lib="NCI_CGAP_Gas4"
		/tissue_type="poorly differentiated adenocarcinoma with signet ring cell features"
		/lab_host="DH10B"
		/note="organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"
BASE COUNT	95 a	70 c 65 g 135 t
ORIGIN		3 others

[illegible]

```

mus_musculatus; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 378)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterson,K.
The WASHU-HHMI Mouse EST Project
Unpublished (1996)
On Nov 29, 1993 this sequence version replaced gi:430248.
Contact: Marra M/Mouse EST project
WASHU-HHMI Mouse EST Project
Washington University School of Medicine#
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:351436
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 367.
Location/Qualifiers
1. .378
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:576848"
/clone_lib="Soares_thymus_2NBMt"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"

```

/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTACCACTCAAGTGGAGCGCGCGGTTTATTTTATTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaldo."

BASE COUNT 116 a 91 c 81 g 90 t
ORIGIN

Query Match 38.4%; Score 28.8; DB 28; Length 378;
Best Local Similarity 75.08; Pred. No. 4.9e+02;
Matches 36; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 10 TTGAATTTTATTTTATTTTGGATATAAGCGCGCGCATGGGC 57
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 64 TTTAATCTTATTTTATTTTGGATATAAGCGCGCGCATGGGC 17

RESULT 14
A1683617/c 423 bp mRNA EST 15-DEC-1999
LOCUS tw52g05.x1 NCI_CGAP_Utl1 Homo sapiens cDNA clone IMAGE:2263352 3'
DEFINITION similar to contains element MBR22 repetitive element ;, mRNA
sequence.
ACCESSION A1683617.1 GI:4893799
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 423)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On May 18, 1998 this sequence version replaced gi:3138026.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html

Insert Length: 1888 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 417.
Location/Qualifiers
FEATURES source
1..423
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2263352"
/clone_lib="NCI_CGAP_Utl1"
/tissue_types="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPOrt5; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 Kb. Life Technologies catalog #:
11538-014"

BASE COUNT 101 a 105 c 78 g 139 t
ORIGIN

Query Match 38.4%; Score 28.8; DB 53; Length 423;
Best Local Similarity 65.68; Pred. No. 4.7e+02;
Matches 42; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 5 AAAAATTGAAATTTATTTTATTTTGGATATAAGCGCGCGCATGGCCGCGG 64
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 151 AAAAGGGGAATTTTCTCTTTTGGCCCAAGGGCGCGCGGTGGGGG 92

BASE COUNT 65 CCAA 68
ORIGIN 91 GCTA 88

RESULT 15
W30577/c 425 bp mRNA EST 11-SEP-1996
LOCUS mc28a08.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
DEFINITION IMAGE:349814 5' similar to gb:U08378 Mus musculus acute phase
response factor (MOUSE);, mRNA sequence.
ACCESSION W30577.1 GI:1310546
VERSION EST.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 425)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

Unpublished (1996)
On Apr 14, 1993 this sequence version replaced gi:692920.
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:221614
Seq primer: ETPPrimer
High quality sequence stop: 423.
Location/Qualifiers
FEATURES source
1..425
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:349814"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Vector: pT7T3D (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTACCACTCAAGTGGAGCGCGCGCATTTTATTTTATTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru KO (Wayne State University)."

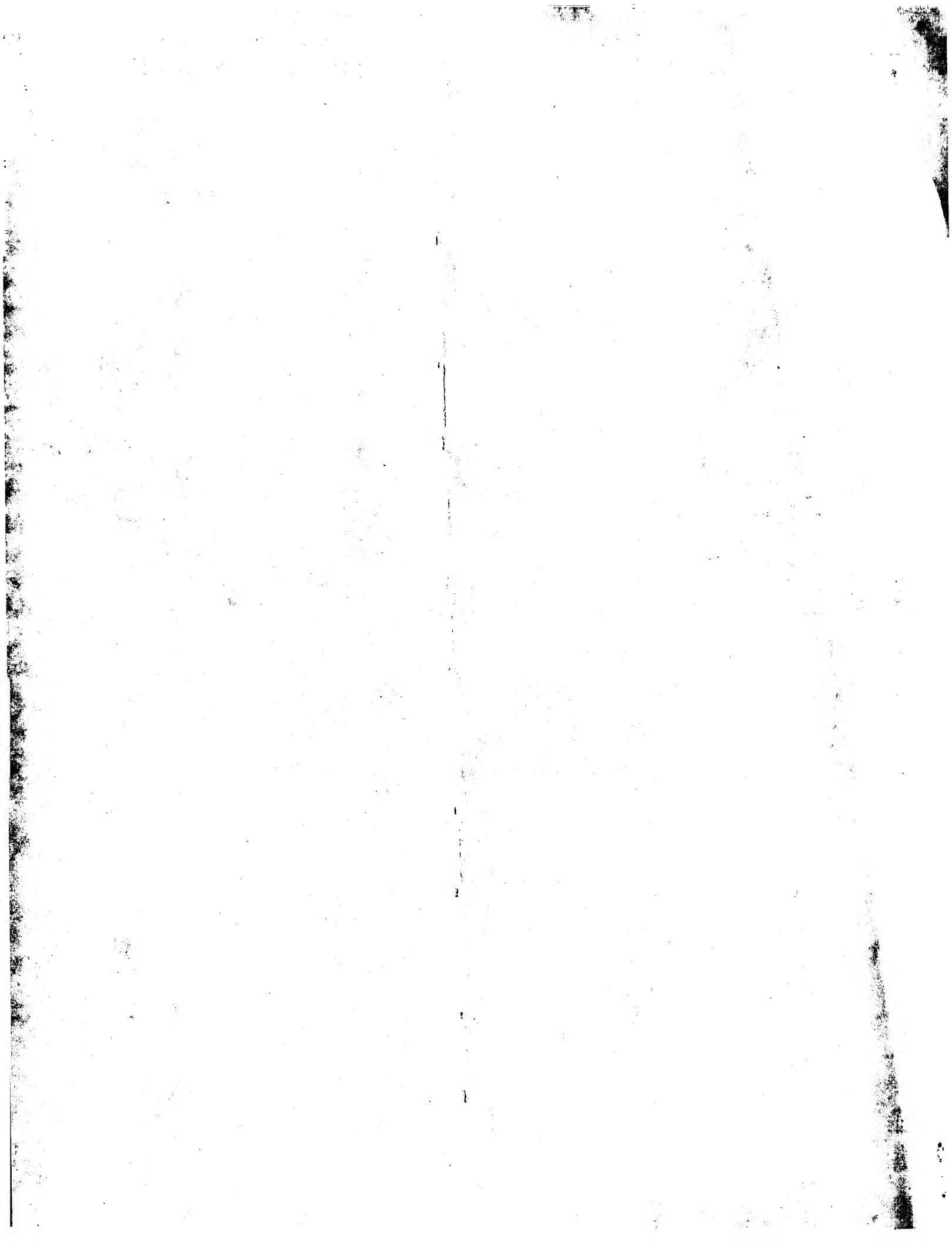
BASE COUNT 125 a 106 c 91 g 103 t
ORIGIN

Query Match 38.4%; Score 28.8; DB 26; Length 425;

Best Local Similarity 75.0%; Pred. No. 4.7e+02;
Matches 36; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 10 TTGAAATTTTATTTTTTTTGGAAATATAAAGCGCGCCGCCATGGGC 57
 ||| | ||||| ||||| ||| ||||| ||| |||||
Db 91 TTTAATTCCTATTTTTTTTTTTTAAATTTAAAGAGGAACCTCTGGGC 44

Search completed: May 29, 2000, 21:13:13
Job time: 36512 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: May 29, 2000, 21:35:19 ; Search time 1214.87 Seconds
(without alignments)
909.139 million cell updates/sec

Title: US-08-935-377-6
Perfect score: 145
Sequence: 1 GCCCAAAATGAAACTA.....TTGTTTTGTGGCCCGGCC 145

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 882769 seqs, 3808571567 residues
Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_bai.*
2: gb_ba2.*
3: gb_om.*
4: gb_ov.*
5: gb_pat.*
6: gb_ph.*
7: gb_pil.*
8: gb_pil2.*
9: gb_prl.*
10: gb_pr2.*
11: gb_pr3.*
12: gb_ro.*
13: gb_sts.*
14: gb_sy.*
15: gb_un.*
16: gb_vl.*
17: em_fun.*
18: em_hum1.*
19: em_hum2.*
20: em_in.*
21: em_om.*
22: em_or.*
23: em_ov.*
24: em_pat.*
25: em_ph.*
26: em_pl.*
27: em_ro.*
28: em_sts.*
29: em_sy.*
30: em_un.*
31: em_vl.*
32: gb_htg1.*
33: gb_htg2.*
34: gb_in1.*
35: gb_in2.*
36: em_bai.*
37: em_ba2.*
38: em_hum3.*
39: em_hum4.*
40: gb_pr4.*
41: gb_htg3.*
42: gb_htg4.*
43: gb_htg5.*

44: gb_htg6.*
45: gb_htg7.*
46: em_htg1.*
47: em_htg2.*
48: em_htg3.*
49: em_hum5.*
50: gb_pl3.*
51: gb_pr5.*
52: gb_htg8.*
53: gb_htg9.*
54: gb_htg10.*
55: gb_htg11.*
56: gb_htg12.*
57: gb_htg13.*
58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	70.8	48.8	4326	14	AF072997	AF072997 Cloning v
2	69	47.6	259	5	A44281	A44281 Sequence 9
3	69	47.6	2958	14	ARBLKSM	X52326 Bluescript
4	69	47.6	2958	14	ARBLKSP	X52331 Bluescript
5	69	47.6	2961	14	ARBL2KSM	X52329 Bluescript
6	69	47.6	2961	14	ARBL2KSP	X52327 Bluescript
7	69	47.6	2964	14	SYNBLKSMV	L08784 Bluescribe
8	69	47.6	2964	14	SYNBLKSPV	L08785 Bluescribe
9	69	47.6	3306	14	SYNBPEN66	D85525 Cloning vec
10	69	47.6	4144	14	XXU35131	U35131 Plasmid pBS
11	69	47.6	4267	14	PRS304	U35136 yeast integ
12	69	47.6	4289	14	XXU35136	U35136 Plasmid pBS
13	69	47.6	4373	14	PRS306	U03438 yeast integ
14	69	47.6	4443	14	PRS303	U03435 yeast integ
15	69	47.6	4670	14	ASAJ5326	AJ005326 pGAII(+)
16	69	47.6	4670	14	ASAJ5329	AJ005329 pGAII(-)
17	69	47.6	4707	14	XXU02374	U02374 Cloning vec
18	69	47.6	4768	14	XXU25061	U25061 Cloning vec
19	69	47.6	4783	14	PRS314	U03440 yeast centr
20	69	47.6	4887	14	PRS316	U03442 yeast centr
21	69	47.6	4950	14	XXU25060	U25060 Cloning vec
22	69	47.6	4967	14	PRS313	U03439 yeast centr
23	69	47.6	5144	14	CVU23751	U23751 Cloning vec
24	69	47.6	5187	14	U34887	U34887 Yeast integ
25	69	47.6	5228	14	XXU25059	U25059 Cloning vec
26	69	47.6	5504	14	PRS305	U03437 Yeast integ
27	69	47.6	5634	14	CVU14125	U14125 Cloning vec
28	69	47.6	6018	14	PRS315	U03441 Yeast centr
29	69	47.6	6340	14	ASAJ5323	AJ005323 pCPI(-) K
30	69	47.6	9655	14	SYNPR8V	AR001531 Cloning v
31	68.2	47.0	793	5	AR060142	AR060142 Sequence
32	67.8	46.8	338	1	BSFX91477	X91477 Bacterial s
33	67.8	46.8	4323	14	AF072538	AF072538 Cloning v
34	67.4	46.5	3240	8	AF015771	AF015771 Magnaport
35	67	46.2	144	14	SYNPLKRA	M22847 Cloning vec
36	67	46.2	147	14	SYNPLKRB	M22848 Cloning vec
37	67	46.2	799	4	AU082676	U82676 Aphelocoma
38	67	46.2	2958	14	ARBLKSM	X52324 Bluescript
39	67	46.2	2958	14	ARBLKSP	X52325 Bluescript
40	67	46.2	2961	14	ARBL2SKM	X52330 Bluescript
41	67	46.2	2961	14	ARBL2SKP	X52328 Bluescript
42	67	46.2	2964	14	SYNBLDKPV	L08787 Bluescribe
43	67	46.2	2964	14	SYNBLKSMV	L08786 Bluescribe
44	67	46.2	3062	14	CVU47947	U47947 Cloning vec
45	67	46.2	3228	14	CVE7829	AJ007829 Cloning v

ALIGNMENTS

```

RESULT 1
AF072997 1 AF072997 4326 bp DNA circular SYN 02-JUL-1998
LOCUS Cloning vector pCMVMTAG3a, complete sequence.
DEFINITION AF072997
ACCESSION AF072997
VERSION AF072997.1 GI:3284001
KEYWORDS
SOURCE Cloning vector pCMVMTAG3a.
ORGANISM artificial sequence; vectors.
REFERENCE 1 (bases 1 to 4326)
AUTHORS Lu.Q.
TITLE pCMVMTAG3a
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4326)
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (22-JUN-1998) Technical Services, Stratagene, 11011 N.
Torrey Pines Rd., San Diego, CA 92037, USA
FEATURES
source
Location/Qualifiers
1..4326
/organism="Cloning vector pCMVMTAG3a"
/db_xref="taxon:78249"
/lab_host="Escherichia coli K12"
1..602
/note="CMV promoter"
713..799
/note="multiple cloning site"
859..1242
/note="SV40 polyA signal region"
1242..1703
/note="fl origin of replication"
1704..1835
/note="pBla promoter"
1836..2227
/note="SV40 origin of replication"
2228..3022
/function="kanamycin resistance"
/codon_start=1
/transl_table=11
/product="aminoglycoside phosphotransferase"
/protein_id="AAC25426.1"
/db_xref="GI:3284002"
/transl_table="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGR
PVLVFTDLGSLNELQDEAARLSWLATTCGPAAYLDVVVTEAGRDWLLIGVPGQDL
LSSHAPAEKVSIMADMRRLTLDPATCPDHOAKHRIETARTMEAGLVDDDDLDE
EHQGLAPAEFLARKASMPDGEDLVVTHGDACLPINIMVNGRESGFIDCGRLGVADRY
QDIALATRTAEELGGEWADRFLVLGIAAPDSQRIATFRLDDEF"
rep_origin
3469..4316
/note="ColE1 origin of replication"
BASE COUNT 1056 a 1116 c 1107 g 1047 t
ORIGIN

Query Match 48.8%; Score 70.8; DB 14; Length 4326;
Best Local Similarity 76.3%; Pred. No. 1.9e-12;
Matches 87; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 12 GAATAACATGATCTATTTATTTCACCGCGCGCGTGGATCCCGCGCTCAGGAATTC 71
|||||
DB 681 GCAGAACTCATCTCTGTAAGAGGATCTGCGCGCGCGATCCCGCGCTCAGGAATTC 740
|||||
QY 72 GATATCAAGCTATCGATACCGTGCACCTCGAGGGGGGCGCTAACTAACTAATT 125
|||||
DB 741 GATATCAAGCTATCGATACCGTGCACCTCGAGGGGGGCGCTACCTTAATT 794
|||||

RESULT 2
A44281/c 259 bp DNA PAT 07-MAR-1997
LOCUS Sequence 9 from Patent WO9511300.
DEFINITION

```

```

ACCESSION A44281
VERSION A44281.1 GI:2299115
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 259)
AUTHORS Chandley,A.C., Kun,M., Sharkey,A.M., Hargreave,T.B. and Cooke,H.J.
TITLE AZOOSPERMIA IDENTIFICATION AND TREATMENT
JOURNAL Patent: WO 9511300-A 9 27-APR-1995;
MEDICAL RES COUNCIL (GB)
COMMENT Other publication AU 7947794 950508.
FEATURES
source
Location/Qualifiers
1..259
/organism="unidentified"
/db_xref="taxon:32644"
/clone_lib="MOUSE GENOMIC LIBRARY"
/clone="M3.2"
BASE COUNT 67 a 58 c 63 g 69 t 2 others
ORIGIN

Query Match 47.6%; Score 69; DB 5; Length 259;
Best Local Similarity 88.2%; Pred. No. 5.7e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 105
|||||
DB 112 GTGGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 53
|||||
QY 106 GGGGGCTTAACCTAACTAATTGTT 130
|||||
DB 52 GGGGGCTTAACCTAACCAGCTTTGTT 28
|||||

RESULT 3
ARBLKSM
LOCUS ARBLKSM 2958 bp DNA circular SYN 10-MAY-1995
DEFINITION pBluescript KS(-) vector DNA, phagemid excised from lambda ZAP.
ACCESSION X52326
VERSION X52326.1 GI:58064
KEYWORDS artificial sequence; cloning vector; expression vector; vector.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 2958)
AUTHORS Thomas,E.A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
11099 North Torrey Pines Rd., La Jolla, CA 92037, USA
REFERENCE 2 (bases 1 to 2958)
AUTHORS Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
TITLE Lambda ZAP: a bacteriophage lambda expression vector with in vivo
excision properties
JOURNAL Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE 88319944
REFERENCE 3 (bases 1 to 2958)
AUTHORS Altting-Mees,M.A. and Short,J.M.
TITLE pBluescript II: gene mapping vectors
JOURNAL Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE 90067967
FEATURES
source
Location/Qualifiers
1..2958
/organism="synthetic construct"
/db_xref="taxon:32630"
1..2958
/note="phagemid pBluescript KS(-)"
BASE COUNT 708 a 754 c 731 g 765 t
ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 2958;
Best Local Similarity 88.2%; Pred. No. 7e-12;

```

```
Matches 75: Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
|||||
Db 687 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746
|||||
QY 106 GGGGGCCTAACTAACTAATTTGTT 130
|||||
Db 747 GGGGGCCCGGTACCCAGCTTTGTT 771
|||||

RESULT 4
ARBLKSP
LOCUS ARBLKSP 2958 bp DNA circular SYN 11-JUN-1998
DEFINITION pBluescript KS(+) vector DNA, phagemid excised from lambda ZAP.
ACCESSION X52331
VERSION X52331.1 GI:58065
KEYWORDS artificial sequence; cloning vector; expression vector; vector.
SOURCE synthetic construct.
ORGANISM
AUTHORS Thomas,E.A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
11099 North Torney Pines Rd., La Jolla, CA 92037, USA
AUTHORS
REFERENCE 2 (bases 1 to 2958)
SHORT,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
LAMBDA ZAP: a bacteriophage lambda expression vector with in vivo
excision properties
Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE 88319944
REFERENCE 3 (bases 1 to 2961)
AUTHORS Altling-Mees,M.A. and Short,J.M.
TITLE pBluescript II: gene mapping vectors
JOURNAL Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE 90067967
FEATURES
source
1..2958
/organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 749 a 734 c 751 g 724 t
BASE COUNT 749 a 734 c 751 g 724 t
ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 2958;
Best Local Similarity 88.2%; Pred. No. 7e-12; Indels 0; Gaps 0;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
|||||
Db 687 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746
|||||
QY 106 GGGGGCCTAACTAACTAATTTGTT 130
|||||
Db 747 GGGGGCCCGGTACCCAGCTTTGTT 771
|||||

RESULT 5
ARBL2KSM
LOCUS ARBL2KSM 2961 bp DNA circular SYN 10-MAY-1995
DEFINITION pBluescript II KS(-) vector DNA, phagemid excised from lambda
ZAPII.
ACCESSION X52329
VERSION X52329.1 GI:58060

Query Match 47.6%; Score 69; DB 14; Length 2958;
Best Local Similarity 88.2%; Pred. No. 7e-12; Indels 0; Gaps 0;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
|||||
Db 687 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746
|||||
QY 106 GGGGGCCTAACTAACTAATTTGTT 130
|||||
Db 747 GGGGGCCCGGTACCCAGCTTTGTT 771
|||||

RESULT 6
ARBL2KSP
LOCUS ARBL2KSP 2961 bp DNA circular SYN 10-MAY-1995
DEFINITION pBluescript II KS(+) vector DNA, phagemid excised from lambda
ZAPII.
ACCESSION X52327
VERSION X52327.1 GI:58061
KEYWORDS artificial sequence; cloning vector; expression vector; vector.
SOURCE synthetic construct.
ORGANISM
AUTHORS Thomas,E.A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
11099 North Torney Pines Rd., La Jolla, CA 92037, USA
AUTHORS
REFERENCE 2 (bases 1 to 2961)
SHORT,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
LAMBDA ZAP: a bacteriophage lambda expression vector with in vivo
excision properties
Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE 88319944
REFERENCE 3 (bases 1 to 2961)
AUTHORS Altling-Mees,M.A. and Short,J.M.
TITLE pBluescript II: gene mapping vectors
JOURNAL Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE 90067967
FEATURES
source
1..2961
/organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 706 a 758 c 735 g 762 t
BASE COUNT 706 a 758 c 735 g 762 t
ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 2961;
Best Local Similarity 88.2%; Pred. No. 7e-12; Indels 0; Gaps 0;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
|||||
Db 687 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746
|||||
QY 106 GGGGGCCTAACTAACTAATTTGTT 130
|||||
Db 747 GGGGGCCCGGTACCCAGCTTTGTT 771
|||||
```

```
KEYWORDS artificial sequence; cloning vector; expression vector; vector.
SOURCE synthetic construct.
ORGANISM
REFERENCE 1 (bases 1 to 2961)
AUTHORS Thomas,E.A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
11099 North Torney Pines Rd., La Jolla, CA 92037, USA
AUTHORS
REFERENCE 2 (bases 1 to 2961)
SHORT,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
LAMBDA ZAP: a bacteriophage lambda expression vector with in vivo
excision properties
Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE 88319944
REFERENCE 3 (bases 1 to 2961)
AUTHORS Altling-Mees,M.A. and Short,J.M.
TITLE pBluescript II: gene mapping vectors
JOURNAL Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE 90067967
FEATURES
source
1..2961
/organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 706 a 758 c 735 g 762 t
BASE COUNT 706 a 758 c 735 g 762 t
ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 2961;
Best Local Similarity 88.2%; Pred. No. 7e-12; Indels 0; Gaps 0;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
|||||
Db 687 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746
|||||
QY 106 GGGGGCCTAACTAACTAATTTGTT 130
|||||
Db 747 GGGGGCCCGGTACCCAGCTTTGTT 771
|||||

RESULT 6
ARBL2KSP
LOCUS ARBL2KSP 2961 bp DNA circular SYN 10-MAY-1995
DEFINITION pBluescript II KS(+) vector DNA, phagemid excised from lambda
ZAPII.
ACCESSION X52327
VERSION X52327.1 GI:58061
KEYWORDS artificial sequence; cloning vector; expression vector; vector.
SOURCE synthetic construct.
ORGANISM
AUTHORS Thomas,E.A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
11099 North Torney Pines Rd., La Jolla, CA 92037, USA
AUTHORS
REFERENCE 2 (bases 1 to 2961)
SHORT,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
LAMBDA ZAP: a bacteriophage lambda expression vector with in vivo
excision properties
Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE 88319944
REFERENCE 3 (bases 1 to 2961)
AUTHORS Altling-Mees,M.A. and Short,J.M.
TITLE pBluescript II: gene mapping vectors
JOURNAL Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE 90067967
FEATURES
source
1..2961
/organism="synthetic construct"
```

Wed May 31 10:04:53 2000

```

/db_xref="taxon:32630"
misc_feature      1..2961
                  /note="phagemid pBluescriptII KS(+)"
BASE COUNT       747 a 738 c 755 g 721 t
ORIGIN
Query Match      47.6%; Score 69; DB 14; Length 2961;
Best Local Similarity 88.2%; Pred. No. 7e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGGCTGCAGGAATTCGATATCGATACGCTATCGATACCGTCGACCTCGAGG 105
      |||||
DB 687 GTGGATCCCCGGGCTGCAGGAATTCGATATCGATACGCTATCGATACCGTCGACCTCGAGG 746
      |||||

QY 106 GGGGGCCCTAACTAACTAATTTTGT 130
      |||||
DB 747 GGGGGCCCGGTACCCAGCTTTTGT 771
      |||||

RESULT 7
SYNBLKSMV      2964 bp      DNA circular      SYN      26-JUL-1993
LOCUS      Bluescribe KS Minus cloning vector.
DEFINITION      L08784
ACCESSION      L08784.1 GI:310728
VERSION
KEYWORDS
SOURCE      Synthetic construct DNA.
ORGANISM      artificial construct.
REFERENCE      1 (bases 1 to 2964)
AUTHORS      Gilbert, W.
TITLE      Obtained from VecBase 3.0
JOURNAL      Unpublished (1991)
COMMENT      These data and their annotation were supplied to GenBank by Will
      Gilbert under the auspices of the Genbank Curator Program.
      Bluescribe KS Minus - Cloning vector
ENTRY BLUEKSM #TYPE DNA CIRCULAR
TITLE Bluescribe KS Minus - Cloning vector
DATE 28-JAN-1987
#sequence 02-FEB-1987
#sequence 04-MAR-1987
#sequence 03-APR-1987
ACCESSION VB0077
SOURCE artificial
REFERENCE
      #number 1
      #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
      J.
      #journal Gene (1987) in press
      #citation Sequence data from StrataGene
      #comment sequence correction according to StrataGene COMMENT
      Obtained from StrataGene on floppy disc.
      Revised 02-FEB-1987 by F. Pfeiffer:
      1409/10 'Aa' to 'TA' to match revised sequence of pBR322
      Revised 4-MAR-1987 to match sequence of pUC19 on request
      of StrataGene
      Polylinker region revised 03-APR-1987 according to StrataGene
      COMMENT
      The stand shown corresponds to pUC19c.
      As in the published sequence of pUC19c, The M13mp19 lacZ region
      is on the complementary strand.
      COMMENT
      This vector contains the fl origin so that the minus strand
      can be obtained upon fl superinfection.
      KEYWORDS
      CROSSREFERENCE
      #parent
      VecBase(3):BlueM13m
      #parent
      VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
      VecSource(3):bgalks, GenBank(50):PFI
      #brother

```



```

#sequence 02-FEB-1987
#sequence 04-MAR-1987
#sequence 03-APR-1987
ACCESSION VB0078
SOURCE artificial
REFERENCE
#number 1
#authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
J.
#journal Gene (1987) in press
#citation sequence data from Stratagene
#comment sequence correction according to Stratagene COMMENT
Obtained from StrataGene on floppy disc.
Revised 02-FEB-1987 by F. Pfeiffer:
1409/10 'Ar' to 'TA' to match revised sequence of PBR322
Revised 4-MAR-1987 to match sequence of pUC19 on request
of Stratagene
Polylinker region revised 03-APR-1987 according to Stratagene
COMMENT
The stand shown corresponds to pUC19c.
As in the published sequence of pUC19c, The M13mp19 lacZ region
is on the complementary strand.
COMMENT
This vector contains the fl origin so that the plus strand
can be obtained upon fl superinfection.
KEYWORDS
CROSSREFERENCE
#parent
VecBase(3):BlueM13p
#parent
VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
VecSource(3):bGalKS, GenBank(50):PFI
#brother
VecBase(3):BlueKSm, VecBase(3):BlueSKp
PARENT
Features of BlueKSp (2964 bp)
residue source
3- 458 5943-5488 (c) phage f1
460- 624 236- 400 pUC19c
626- 645 1- 20 T7 promoter
653- 760 108- 1 (c) BlueKS-polylinker
772- 791 20- 1 (c) T3 promoter
795-2964 448-2617 pUC19c
Conflict (cfl) and Mutations (mut): none
PARENT
Features of BlueKSp (2964 bp)
residue source
3- 458 5943-5488 (c) phage f1
460- 624 449- 285 (c) pUC19
626- 645 1- 20 T7 promoter
653- 760 108- 1 (c) BlueKS-polylinker
772- 791 20- 1 (c) T3 promoter
795-1031 237- 1 (c) pUC19
1032-2964 2686- 754 (c) pUC19
Conflict (cfl) and Mutations (mut): none
FEATURE
643 start of T7-RNA synthesis
774 (c) start of T3-RNA synthesis
1976-2764 789-1 (c) Ap-R; b-lactamase
POLYLINKER
KpnI-DraIII-ApaI-XhoI-SalI-HindIII-EcoRV-EcoRI-PstI-
SmaI-BamHI-SpeI-XbaI-NotI-XmaII-BstXI-SacII-SacI SELECTION
#resistance Ap
#indicator beta-galactosidase
SUMMARY BlueKSp #length 2964 #checksum 690.
Location/Qualifiers
1..2964
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 750 a 736 c 755 g 723 t
ORIGIN

#sequence 02-FEB-1987
#sequence 04-MAR-1987
#sequence 03-APR-1987
ACCESSION VB0078
SOURCE artificial
REFERENCE
#number 1
#authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
J.
#journal Gene (1987) in press
#citation sequence data from Stratagene
#comment sequence correction according to Stratagene COMMENT
Obtained from StrataGene on floppy disc.
Revised 02-FEB-1987 by F. Pfeiffer:
1409/10 'Ar' to 'TA' to match revised sequence of PBR322
Revised 4-MAR-1987 to match sequence of pUC19 on request
of Stratagene
Polylinker region revised 03-APR-1987 according to Stratagene
COMMENT
The stand shown corresponds to pUC19c.
As in the published sequence of pUC19c, The M13mp19 lacZ region
is on the complementary strand.
COMMENT
This vector contains the fl origin so that the plus strand
can be obtained upon fl superinfection.
KEYWORDS
CROSSREFERENCE
#parent
VecBase(3):BlueM13p
#parent
VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
VecSource(3):bGalKS, GenBank(50):PFI
#brother
VecBase(3):BlueKSm, VecBase(3):BlueSKp
PARENT
Features of BlueKSp (2964 bp)
residue source
3- 458 5943-5488 (c) phage f1
460- 624 236- 400 pUC19c
626- 645 1- 20 T7 promoter
653- 760 108- 1 (c) BlueKS-polylinker
772- 791 20- 1 (c) T3 promoter
795-2964 448-2617 pUC19c
Conflict (cfl) and Mutations (mut): none
PARENT
Features of BlueKSp (2964 bp)
residue source
3- 458 5943-5488 (c) phage f1
460- 624 449- 285 (c) pUC19
626- 645 1- 20 T7 promoter
653- 760 108- 1 (c) BlueKS-polylinker
772- 791 20- 1 (c) T3 promoter
795-1031 237- 1 (c) pUC19
1032-2964 2686- 754 (c) pUC19
Conflict (cfl) and Mutations (mut): none
FEATURE
643 start of T7-RNA synthesis
774 (c) start of T3-RNA synthesis
1976-2764 789-1 (c) Ap-R; b-lactamase
POLYLINKER
KpnI-DraIII-ApaI-XhoI-SalI-HindIII-EcoRV-EcoRI-PstI-
SmaI-BamHI-SpeI-XbaI-NotI-XmaII-BstXI-SacII-SacI SELECTION
#resistance Ap
#indicator beta-galactosidase
SUMMARY BlueKSp #length 2964 #checksum 690.
Location/Qualifiers
1..2964
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 750 a 736 c 755 g 723 t
ORIGIN

Query Match 47.68; Score 69; DB 14; Length 2964;
Best Local Similarity 88.28; Pred. No. 7e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 46 GTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
|||||
Db 687 GTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746
Qy 106 GGGGCCCTAACTAATAATTTGTT 130
|||||
Db 747 GGGGCCCGGTACCGAGCTTTGTT 771

RESULT 9
SYNPBEN66/c 3306 bp DNA circular SYN 06-FEB-1999
LOCUS
DEFINITION Cloning vector pBEN66 DNA for aminoglycoside 3'-phosphotransferase,
beta-lactamase, complete cds.
ACCESSION D85525
VERSION D85525.1 GI:1345433
KEYWORDS plasmid; aminoglycoside 3'-phosphotransferase; beta-lactamase.
SOURCE Cloning vector pBEN66 (lab_host:E.coli) plasmid:pBEN66 DNA.
ORGANISM Cloning vector pBEN66
artificial sequence; vectors.
REFERENCE 1 (bases 1 to 3306)
AUTHORS Yamamoto,Y.
TITLE Direct Submission
JOURNAL
PUBLISHED 23-MAY-1996 to the DDBJ/EMBL/GenBank databases.
Yoshihiro Yamamoto, Hyogo College of Medicine, Department of
Genetics, Mukogawa-cho 1-1, Nishinomiya, Hyogo 663, Japan
(Tel:0798-45-6587, Fax:0798-40-7639)
REFERENCE 2 (sites)
AUTHORS Yamamoto,Y. and Furuyama,J.
TITLE One-step disruption by circular DNA in Escherichia coli
JOURNAL Unpublished (1996)
FEATURES
Location/Qualifiers
1..3306
/organism="Cloning vector pBEN66"
/plasmid="pBEN66"
/db_xref="taxon:47800"
/lab_host="E.coli"
19..36
/promoter
/feature="T3 promoter"
260..1075
/gene="kan"
260..1075
/gene="kan"
/feature="derived from Tn903; kanamycin resistance gene"
/citation=[2]
/codon_start=1
/transl_table=11
/product="aminoglycoside 3'-phosphotransferase"
/protein_id="BAAL2824.1"
/db_xref="GI:1345434"
/translation="MSHQRETSCSRPLNSMDADLYGYKWARDNVQSGATYRLY
GKPAPELFLKHGSGVANDVDVEMRLNLTETEMPLTIRKHFRTDDAWLLTAIP
GKPAQVLEETPDGSEINVDALAVFLRLHSIPVCNPFNSDRVFRQAQSRMNGL
VDSDFDDEGRNVEQVWKEMHLLPFSVDVYTHGDFSLDNLIFDEKGLIGDIVG
RVGIADYQDLAILWNLGFEFSPLQKRLFKYKIDNPDMNKLQFHLMLDEFF"
complement(2209..3069)
/gene="amp"
complement(2209..3069)
/gene="amp"
/feature="ampicillin resistance gene"
/citation=[1]
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="BAAL2825.1"
/db_xref="GI:1345435"
/translation="MSIQHFRVALIPFFAAFCPLPVFAHPETLVKVAEDOLGARVGY
IELDLSKILLESFRRPERPMSTMTEFKVLGGAVLSRIDAGQEOIGRIHYSONDLVE
YSPVTERHLDGNTVRELCSNAITMSDNTAANLLLTIGGPKELTAFLLHMGDHTRL"

```

BASE COUNT	1084 a	1009 c	959 g	1092 t
------------	--------	--------	-------	--------

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 4289)	Alexeyev, M.F., Shokolenco, I.N. and Croughan, T.P.	Improved antibiotic-resistance gene cassettes and

BASE COUNT	1084 a	1009 c	959 g	1092 t
------------	--------	--------	-------	--------

for Escherichia coli vector construction and in vitro
deletion/insertion mutagenesis

JOURNAL
MEDLINE

Gene 160 (1), 63-67 (1995)
95354958

REFERENCE
AUTHORS

2 (bases 1 to 4289)
Hengen, P.N.

JOURNAL
TITLE

Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical
Biology, National Cancer Institute, Frederick, MD 21702-1201, USA

FEATURES
source

1, 4289
Location/Qualifiers

/organism="Cloning vector pBSL97"

/plasmid="pBSL97"

/db_xref="taxon:42709"

complement(804..1598)

/codon_start=1

/transl_table=11

/product="neomycin phosphotransferase"

/protein_id="AAC53629.1"

/db_xref="GI:984924"

/translation="MIEQDGLHAGSPAAMVERLFGYDWAQQTIGCSDAVFLRLSAQGR

PVLFTDLGALNEQDEARLHSLWLTGVPAAVLVDVVTAGRDWLLLGVEVPGQDL

LSSHAPAEKYSIMADAMRHLTLPATCPFDHQAHRERARTRMEAGLVDDDDLDE

EHQGLAPAELEFARLPDCEDLVYTHGDACLPLNIMVENGFSGFIDCGRLGVADRY

QDIALADIAEELGGEWADFLVLGIAAPDSQRIAFYRLLEFF"

complement(3301..4161)

/EC_number="3.5.2.6"

/codon_start=1

/transl_table=11

/function="ampicillin resistance"

/product="beta-lactamase"

/protein_id="AAC53630.1"

/db_xref="GI:984925"

/translation="MSIQHFRVALIPFFAAFLPVAHPETLVKVKDAEDQLGARVGY

IELDLSGKILLESFPERFPMWTFKLLCGAVLSRIDAGQEOGLRIHYSONDLVE

YSPVTEKLTGCMVRELCSAITMSDNTANLLLTIGGPKELTAFLHNMGDHYTRL

DRWPELNEATPNDERTMPVAMATTLKLLTGELLILASRQQLIDWNEADKVAQPL

LRSLAPAGWFAADKSGAGERSGRIIAALGPDGKPSRIVIITGSGQATMDERNKQIA

EIGASLIKHW"

BASE COUNT 1024 a 1157 c 1111 g 997 t

ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 4289;
Best Local Similarity 88.2%; Pred. No. 7.2e-12;

Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105

Db 2015 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2074

QY 106 GGGGGCCTAACTAACTAATTTGTT 130

Db 2075 GGGGGCCCGGTACCCAGCTTTGTT 2099

FEATURES

source

1, 4443
Location/Qualifiers

/organism="Cloning vector pBS303"

/db_xref="taxon:31826"

BASE COUNT 1149 a 1048 c 1109 g 1137 t

ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 4443;
Best Local Similarity 88.2%; Pred. No. 7.2e-12;

Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105

Db 2107 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2166

QY 106 GGGGGCCTAACTAACTAATTTGTT 130

Db 2167 GGGGGCCCGGTACCCAGCTTTGTT 2191

FEATURES

source

1, 4443
Location/Qualifiers

/organism="Cloning vector pBS303"

/db_xref="taxon:31826"

BASE COUNT 1149 a 1048 c 1109 g 1137 t

ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 4443;
Best Local Similarity 88.2%; Pred. No. 7.2e-12;

Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

REFERENCE
AUTHORS

2 (bases 1 to 4373)
Stillman, D.J.

JOURNAL
TITLE

Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
and Molecular Biology, University of Utah Medical Center, Salt Lake
City, UT 84132, USA

FEATURES
source

1, 4373
Location/Qualifiers

/organism="Cloning vector pBS306"

/db_xref="taxon:31829"

BASE COUNT 1181 a 987 c 1106 g 1099 t

ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 4373;
Best Local Similarity 88.2%; Pred. No. 7.2e-12;

Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105

Db 2027 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2086

QY 106 GGGGGCCTAACTAACTAATTTGTT 130

Db 2087 GGGGGCCCGGTACCCAGCTTTGTT 2111

FEATURES

source

1, 4373
Location/Qualifiers

/organism="Cloning vector pBS306"

/db_xref="taxon:31829"

BASE COUNT 1181 a 987 c 1106 g 1099 t

ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 4373;
Best Local Similarity 88.2%; Pred. No. 7.2e-12;

Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105

Db 2027 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2086

QY 106 GGGGGCCTAACTAACTAATTTGTT 130

Db 2087 GGGGGCCCGGTACCCAGCTTTGTT 2111

FEATURES

source

1, 4373
Location/Qualifiers

/organism="Cloning vector pBS306"

/db_xref="taxon:31829"

BASE COUNT 1181 a 987 c 1106 g 1099 t

ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 4373;
Best Local Similarity 88.2%; Pred. No. 7.2e-12;

Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105

Db 2107 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2166

QY 106 GGGGGCCTAACTAACTAATTTGTT 130

Db 2167 GGGGGCCCGGTACCCAGCTTTGTT 2191

FEATURES

source

1, 4373
Location/Qualifiers

/organism="Cloning vector pBS306"

/db_xref="taxon:31829"

BASE COUNT 1181 a 987 c 1106 g 1099 t

ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 4373;
Best Local Similarity 88.2%; Pred. No. 7.2e-12;

Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105

Db 2107 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2166

QY 106 GGGGGCCTAACTAACTAATTTGTT 130

Db 2167 GGGGGCCCGGTACCCAGCTTTGTT 2191

FEATURES

source

1, 4373
Location/Qualifiers

/organism="Cloning vector pBS306"

/db_xref="taxon:31829"

BASE COUNT 1181 a 987 c 1106 g 1099 t

ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 4373;
Best Local Similarity 88.2%; Pred. No. 7.2e-12;

Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Wed May 31 10:04:53 2000

LOCUS ASAJ5326 4670 bp DNA circular SYN 08-FEB-1999
DEFINITION pgAII(+) KS positive selection cloning vector gltS gene.
ACCESSION AJ005326
VERSION AJ005326.1 GI:4028984
KEYWORDS gltS gene; glutamate permease.
SOURCE synthetic construct.
ORGANISM synthetic construct
1 (bases 1 to 4670)
REFERENCE Gal, J.
AUTHORS Direct Submission
TITLE Submitted (03-APR-1998) Gal J., Institute for Biotechnology, Bay
JOURNAL Zoitan Foundation for Applied Research, Szeged, Derkovits fasor 2.,
MEDLINE 6726, HUNGARY
FEATURES
1 (bases 1 to 4670)
2 (bases 1 to 4670)
Gal, J., Szekeres, S., Schnell, R., Pongor, S., Simoncsits, A. and
Kalman, M.
A positive selection cloning system based on the gltS gene of
Escherichia coli
Anal. Biochem. 266 (2), 235-238 (1999)
99107575
Location/Qualifiers
1..4670
/organism="synthetic construct"
/db_xref="taxon:32630"
852..2231
/gene="gltS"
852..2231
/gene="gltS"
/codon_start=1
/transl_table=11
/product="glutamate permease"
/protein_id="CAA06473.1"
/db_xref="GI:4028985"
/translation="MFHDTLTLVAATLLGLGRKLVHVSFLKYYTIPVPVAGLL
VALALLVKKMGWNEFDMSLRDPMLAFAFTICLNANTASLRAGRVVGIELIVV
GILVMQNAIGIGMSALLGLDPLMGLLGSITLGGHGTGAWSKLFIERYGFTNAIEV
AMACATFGVLGGLGEPVQRLTIGKNSWVFPSPSTVSISLNSCSPDPLV
LERPPRWSSNPYSEYARYLVKHSTTPNGIPDDQEVPTAFKPDVGRMTSLVLI
ETIALICLTAVKVAOLLAGTAFELPFCVLFVGVLSNGLSIMGFYRFRVAVS
VLGNVSLSLFLAMALGKLMLASLALPMLAILVQTIEMALYAIFFVTRMMGKNYD
AAVLAAGHCGFGLGATPTAIAKMQAITEFGFSPHMAFLVVPVMGAFIDIVNALVIKL
YLMFLPIFAG"

BASE COUNT 1040 a 1165 c 1231 g 1234 t
ORIGIN

Query Match 47.6%; Score 69; DB 14; Length 4670;
Best Local Similarity 88.2%; Pred. No. 7.2e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCGGGCTCGAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
|||||
Db 1500 GTGGATCCCGGGCTCGAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 1441
|||||
QY 106 GGGGGCCTAACTAACTTTTGT 130
|||||
Db 1440 GGGGGCCTAACTAACTTTTGT 1416
|||||

Search completed: May 29, 2000, 21:35:29
Job time: 36861 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:58:11 ; Search time 1446.77 Seconds
(without alignments)
25.075 Million cell updates/sec

Title: US-08-935-377-6

Perfect score: 145
Sequencing: 1 GGCAGAAATGAAACTA.....TTGTTTTTGGGCGCGGCC 145

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	47.6	259	1 Q87664	Mouse azoospermia
2	69	47.6	501	1 T04866	Nucleotide analogu
3	67	46.2	551	1 X20513	Polynucleotide seq
4	67	46.2	752	1 X31294	E. coli J96 pathog
5	67	46.2	1091	1 T75006	Human endogenous r
6	67	46.2	1122	1 T75005	Human endogenous r
7	67	46.2	3792	1 Q48463	Plasmid pg+host4 c
8	67	46.2	4226	1 T39485	Human steroidogene
9	67	46.2	5234	1 Q48464	Plasmid pg+host5 c
10	67	46.2	6722	1 Q48465	Plasmid pg+host6 c
11	67	46.2	10529	1 V09028	Maize fluory2 gene
12	67	46.2	12814	1 X24730	Swedish-FAD APP ta
13	67	46.2	15692	1 X24731	London-FAD APP tar
14	67	46.2	15692	1 X24732	Swedish/London-FAD
15	67	46.2	15701	1 X24733	Swedish-FAD APP713
16	66.8	46.1	545	1 V68808	Human endogenous r
17	65.4	45.1	2640	1 Q28664	bDAT. cDNA encodin
18	65.4	45.1	5356	1 T43794	Plasmid priPHAT (r
19	65	44.8	84	1 X02800	E. coli biotin DNA
20	65	44.8	3465	1 X02814	DE19731274 Seq ID
21	65	44.8	3481	1 X02815	DE19731274 Seq ID
22	65	44.8	3794	1 X02812	DE19731274 Seq ID
23	65	44.8	3810	1 X02813	DE19731274 Seq ID
24	65	44.8	6596	1 V57377	Maize female-prefe
25	63.6	43.9	545	1 T75010	Human endogenous r
26	62.6	43.2	2150	1 T91037	Yeast checkpoint c
27	62.6	43.2	2150	1 X01271	Yeast RAD17 coding
28	62.4	43.0	685	1 V10190	Stealth virus nucl
29	62.4	43.0	685	1 V12003	Stealth virus plas
30	60.2	41.5	1949	1 X05602	Nucleotide sequenc
31	59	40.7	4164	1 T04575	Plasmid pat-1 sequ
32	59	40.7	4164	1 V22271	PAT-1 (pSD544). DN
33	59	40.7	4164	1 V69740	Nucleotide sequenc

ALIGNMENTS

RESULT 1

Q87664/c
ID Q87664 standard; DNA; 259 BP.
AC Q87664;
DT 06-NAR-1996 (first entry)
DE Mouse azoospermia factor (AZF) gene partial clone M3.2.
KW Azoospermia factor; AZF; male infertility; YRRM gene;
KW Y-chromosome; RNA recognition motif; ds.
OS Mus sp.
FH Key Location/Qualifiers
FT misc_difference 174
FT /*tag= a
/*note= "base n at position 174 is not identified
in the specification"
FT misc_difference 197
FT /*tag= b
/*note= "base n at position 197 is not identified
in the specification"
PN WO9511300-A2.
PD 27-APR-1995.
PF 24-OCT-1994; G02344.
PR 22-OCT-1993; GB-021857.
PR 07-JUL-1994; GB-013760.
PA (MEDI-) MEDICAL RES COUNCIL.
PI Chandley AC, Cooke HJ, Hargreave TB, Kun M, Sharkey AM;
DR WPI: 95-170221/22.
PT Nucleic acid encoding the human azoospermia factor, and probes and
PT antibodies specific for the sequence and encoded polypeptide - may
PT be used in the clinical diagnosis of male infertility
PS Disclosure; Fig 5; 40pp; English.
CC A mouse gene partial sequence (Q87664) showing homology to the human
CC azoospermia factor (AZF) YRRM gene was obtd. from mouse genomic phage
CC clone M3.2 isolated from a genomic library screened with MYK (YRRM1)
CC cDNA (Q87655). M3.2, cloned in Lambda 2001, was mapped to the short arm
CC of the mouse Y-chromosome. A full cDNA sequence (Q87665) was also
CC obtd.
SQ Sequence 259 BP; 67 A; 58 C; 63 G; 69 T;

Query Match 47.6%; Score 69; DB 1; Length 259;
Best Local Similarity 88.2%; Pred. No. 1.8e-15;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 45 GTGGATCCCCGGGTCGAGGAATCGATATCAAGCTTATCGATACCGTCGACCTCGAG 105

Db 112 GTGGATCCCCGGGTCGAGGAATCGATATCAAGCTTATCGATACCGTCGACCTCGAG 53

QY 106 GGGGGCCTAACTAATTTTGT 130

Db 52 GGGGGCCTAACTAATTTTGT 28

RESULT 2

T04866

ID T04866 standard; DNA; 501 BP.

AC T04866;

pTet-Splice. Nucle
Candida Carhol gen
Clone #4 from muta
E. coli biotin DNA
Plasmid pKS varian
Plasmid pPK13/14 D
Plasmid pPK5/6 DNA
Plasmid pPK9/10 DN
Plasmid pPK7/8 DN
Hybrid vector pSF-
DNA sequence of ex

34 58 40.0 5178 1 T49876
35 57.8 39.9 3198 1 T92702
36 57.8 39.9 3198 1 T92869
37 56.4 38.9 459 1 V57250
c 38 56 38.6 78 1 X02799
39 56 38.6 2973 1 V64254
40 56 38.6 3956 1 V64258
41 56 38.6 4088 1 V64255
42 56 38.6 4102 1 V64257
43 56 38.6 4583 1 V64256
44 55 37.9 6505 1 T13394
45 54 37.2 10504 1 V79503

28-JAN-1996 (first entry)
 Nucleotide analogue treated with calf intestinal alkaline phosphatase.
 Nucleotide analogue; alkaline phosphatase; ss.
 Synthetic.
 Key Location/Qualifiers
 FT misc_feature 1..501
 FT /*tag= a
 FT /note= "std. IUPAC codes used"
 PN WO9524505-A.
 PD 14-SEP-1995.
 PF 07-MAR-1995; G00490.
 PR 08-MAR-1994; EP-301636.
 PA (AMSH) AMERSHAM INT PLC.
 PI Reeve MA, Robinson PS;
 DR WPI; 95-328290/42.
 PT Modification of residual fluorescence labelled nucleotide analogues - to
 prevent migration in electrophoretic sequencing gel and interference with
 base calling of DNA chains
 PS Example; Fig 1; 18pp; English.
 CC The invention concerns the modification of residual fluorescence
 labelled nt analogues to prevent migration in electrophoretic
 sequencing gel and interference with base calling of DNA chains. The
 modification involves the use of a phosphatase enzyme to remove at
 least one 5'-phosphate gp. The nts were prepared using ABI Amplitaq
 dye-terminating kit. The template used was 1 microg. of M13 mp8. The
 primer was M13 Universal primer. Samples were recovered. Some were
 subjected to calf intestinal alkaline phosphatase digestion. This
 includes T04866. T04867 was a control nt, which was not treated with
 alkaline phosphatase. The example shows that alkaline phosphatase
 treatment causes removal of the dye-terminator artefacts and allows
 for accurate base calling with the ABI analysis software.
 CC Sequence 501 BP; 98 A; 136 C; 140 G; 121 T;
 SQ

Query Match 47.6%; Score 69; DB 1; Length 501;
 Best Local Similarity 88.2%; Pred. No. 2.2e-15;
 Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 105
 DB 57 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 116
 QY 106 GGGGGCCCTAACTAACTAATTGGTT 130
 DB 117 GGGGGCCCGGTACCCAGCTTTGGTT 141

RESULT 3
 ID X20513 standard; DNA; 651 BP.
 AC X20513;
 DE 05-MAY-1999 (first entry)
 KW Polynucleotide sequence from the genome of Treponema pallidum.
 KW Treponema pallidum infection; syphilis; Borrelia infection; animal;
 OS enzyme production; ds.
 ON Treponema pallidum.
 PN WO9859034-A2.
 PD 30-DEC-1998.
 PR 24-JUN-1998; U13041.
 PR 24-JUN-1997; US-050667.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Fraser CN;
 DR WPI; 99-081273/07.
 PT New isolated Treponema pallidum nucleic acids - used to develop
 products for the detection, diagnosis, characterisation, prevention
 and therapy of T. pallidum infections, particularly syphilis
 PS Claim 1; Page 257; 1150pp; English.
 CC X20500-21243 represent polynucleotide sequences from the genome of
 Treponema pallidum. The sequences can be used for detection,
 diagnosis, characterisation, prevention and therapy for T. pallidum
 infections, particularly syphilis. They can also be used for detecting
 diseases related to Borrelia infections in animals, and for the
 production of biosynthetic products such as enzymes.

SQ Sequence 651 BP; 180 A; 167 C; 152 G; 149 T;
 Query Match 46.2%; Score 67; DB 1; Length 651;
 Best Local Similarity 100.0%; Pred. No. 1.3e-14;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 105
 DB 72 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 13
 QY 106 GGGGGCC 112
 DB 12 GGGGGCC 6

RESULT 5
 ID T75006/c
 AC T75006; standard; DNA; 1091 BP.
 DE T75006 (first entry)
 DT 06-OCT-1997 (first entry)
 DE Human endogenous retroviral sequence 6.
 KW Breast cancer; tumour; B18Agl; prognosis; diagnosis; vaccine; ss.

SQ Sequence 651 BP; 180 A; 167 C; 152 G; 149 T;
 Query Match 46.2%; Score 67; DB 1; Length 651;
 Best Local Similarity 100.0%; Pred. No. 1.3e-14;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 105
 DB 72 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 13
 QY 106 GGGGGCC 112
 DB 12 GGGGGCC 6

RESULT 5
 ID T75006/c
 AC T75006; standard; DNA; 1091 BP.
 DE T75006 (first entry)
 DT 06-OCT-1997 (first entry)
 DE Human endogenous retroviral sequence 6.
 KW Breast cancer; tumour; B18Agl; prognosis; diagnosis; vaccine; ss.

for detection of mutation(s) of this gene that cause congenital lipid adrenal hyperplasia
Claim 1: Pages 23-25; 89pp; English.
The present sequence encodes the human steroidogenesis acute regulatory protein (hSTAR). The hSTAR gene can be analysed for mutations to detect (e.g. prenatally) genetic defects associated with congenital lipid adrenal hyperplasia (CAH), or its transmission to children. CAH can be treated by protein or gene replacement therapy, which can also be used to prevent or treat hypercholesterolaemia.
A human adrenal cortex cDNA library was screened with a mouse STAR probe to isolate a 1.6 kb insert, including an ORF for a 285 residue protein. When it was cloned into pSPORT and expressed in COS-1 cells cotransfected with pP450scab and pADX, it increased the level of pregnenolone synthesis from cholesterol or 20-alpha-hydroxycholesterol.
Sequence 4226 BP; 940 A; 1132 C; 1144 G; 984 T;
Query Match 46.2%; Score 67; DB 1; Length 4226;
Best Local Similarity 100.0%; Pred. No. 2.4e-14; Mismatches 0; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
Db 4149 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4208
QY 106 GGGGGGCC 112
Db 4209 GGGGGGCC 4215
RESULT 9
Q48464/c
ID Q48464 standard; DNA; 5234 BP.
AC Q48464;
DT 18-MAR-1994 (first entry)
DE Plasmid pg+host5 containing Ts replication system.
KW Temperature sensitive replication; antibiotic resistance marker gene;
KW site-specific recombination; chromosomal integration; inactivation;
KW heterologous gene expression; thermosensitive plasmid; ds.
OS Synthetic.
FT Key Location/Qualifiers
FT misc_feature 36..1496
FT /*tag= a
FT /*standard_name= ORI
FT /*note= "origin of replication from pBR322"
FT misc_feature 2640..4393
FT /*tag= b
FT /*note= "from pgK12 (derived from pWV01)"
FT misc_feature 4384..4786
FT /*tag= c
FT /*note= "from PUB110"
FT misc_feature 4787..5234
FT /*tag= d
FT /*note= "from pSK"
W09318164-A.
PD 16-SEP-1993. F00248.
PF 12-MAR-1993; FR-003034.
PR (INRG) INRA INST NAT RECH AGRONOMIQUE.
PI Gruss A, Maguin E;
DR New bacterial plasmid contg. heat sensitive replication system -
PT and marker gene, opt. capable of chromosomal integration, used to
PT inactivate specific gene or introduce heterologous gene
PS Example 2; Fig 10; 73pp; French.
CC Plasmid pGK12 (Appl. Environ. Microbiol., 48; 726 (1984)) contg. two
CC antibiotic resistance marker genes was subjected to mutagenesis with
CC hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC 445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC site, T7 and T3 promoters and sequencing primer binding sites. The
CC resulting plasmid was designated pVE6004 (or pg+host4 - Q48463).
CC To facilitate cloning in E. coli, the AvarI-EcoRI fragment of pBR322
CC (containing the origin of replication and the ampicillin resistance
CC gene) was inserted into NsiI-cleaved, blunt-ended pg+host4 to give
CC pg+host6 (Q48466). The pBR322 ORI allows maintenance of the plasmid in
CC E. coli at 37 deg.C. while the heat-sensitive ORI allows maintenance
CC at 28 deg.C in gram-positive bacteria.
SQ Sequence 6722 BP; 1961 A; 1278 C; 1503 G; 1980 T;
Query Match 46.2%; Score 67; DB 1; Length 6722;
Best Local Similarity 100.0%; Pred. No. 2.8e-14; Mismatches 0; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
Db 6471 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 6412
QY 106 GGGGGGCC 112
Db 6411 GGGGGGCC 6405

site, T7 and T3 promoters and sequencing primer binding sites. The
resulting plasmid was designated pVE6004 (or pg+host4 - Q48463).
To facilitate cloning in E. coli, the 1.4kb AvarI-AlwNI fragment of
pBR322 (containing the origin of replication) was inserted into
NsiI-cleaved pg+host4 to give pg+host5 (Q48464). The pBR322 ORI
allows maintenance of the plasmid in E. coli at 37 deg.C. while the
heat-sensitive ORI allows maintenance at 28 deg.C in gram-positive
bacteria.
Sequence 5234 BP; 1569 A; 973 C; 1133 G; 1559 T;
Query Match 46.2%; Score 67; DB 1; Length 5234;
Best Local Similarity 100.0%; Pred. No. 2.6e-14; Mismatches 0; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
Db 4983 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4924
QY 106 GGGGGGCC 112
Db 4923 GGGGGGCC 4917
RESULT 10
Q48465/c
ID Q48465 standard; DNA; 6722 BP.
AC Q48465;
DT 18-MAR-1994 (first entry)
DE Plasmid pg+host6 containing Ts replication system.
KW Temperature sensitive replication; antibiotic resistance marker gene;
KW site-specific recombination; chromosomal integration; inactivation;
KW heterologous gene expression; thermosensitive plasmid; ds.
OS Synthetic.
PN W09318164-A.
PD 16-SEP-1993. F00248.
PF 12-MAR-1993; FR-003034.
PR (INRG) INRA INST NAT RECH AGRONOMIQUE.
PI Gruss A, Maguin E;
DR WPI; 93-303478/38.
PT New bacterial plasmid contg. heat sensitive replication system -
PT and marker gene, opt. capable of chromosomal integration, used to
PT inactivate specific gene or introduce heterologous gene
PS Disclosure; Fig 11; 73pp; French.
CC Plasmid pGK12 (Appl. Environ. Microbiol., 48; 726 (1984)) contg. two
CC antibiotic resistance marker genes was subjected to mutagenesis with
CC hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC 445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC site, T7 and T3 promoters and sequencing primer binding sites. The
CC resulting plasmid was designated pVE6004 (or pg+host4 - Q48463).
CC To facilitate cloning in E. coli, the AvarI-EcoRI fragment of pBR322
CC (containing the origin of replication and the ampicillin resistance
CC gene) was inserted into NsiI-cleaved, blunt-ended pg+host4 to give
CC pg+host6 (Q48466). The pBR322 ORI allows maintenance of the plasmid in
CC E. coli at 37 deg.C. while the heat-sensitive ORI allows maintenance
CC at 28 deg.C in gram-positive bacteria.
SQ Sequence 6722 BP; 1961 A; 1278 C; 1503 G; 1980 T;
Query Match 46.2%; Score 67; DB 1; Length 6722;
Best Local Similarity 100.0%; Pred. No. 2.8e-14; Mismatches 0; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
Db 6471 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 6412
QY 106 GGGGGGCC 112
Db 6411 GGGGGGCC 6405


```

RESULT 11
V09028/c
ID V09028 standard; DNA: 10529 BP.
AC V09028;
DE 20-JUL-1998 (first entry)
KW Maize fluory2 gene (fl2).
KW Fluory2; fl2 gene; alpha-zein; signal peptide; transgenic plant;
KW seed; essential amino acid; animal feed; maize; rice; wheat;
KW barley; millet; sorghum; ds.
OS Zea mays.
FH Key Location/Qualifiers
FT misc_feature 1..760
FT /tag= a
FT /note= "vector-derived sequence"
FT promoter 761..3824
FT /tag= b
FT /note= "Claim 14"
FT CDS 3825..4613
FT /tag= c
FT sig_peptide 3825..3887
FT /tag= d
FT mat_peptide 3888..4610
FT /tag= e
FT W09802563-A1.
PN 22-JAN-1998.
PD 11-JUL-1997; U11723.
PF 11-JUL-1997; US-021833.
PR (PTON-) PIONEER HI-BRED INT INC.
PA (UFAR-) UNIV ARIZONA STATE.
PI Beach L, Coleman CE, Larkins BA;
DR WPI: 98-110609/10.
PR P-PSDB: W23977.
DE Cereal plants containing trans-gene expressing fusion that includes
PT signal peptide of the fl2 maize gene - and protein having high
PT content of essential amino acids, producing feeds of improved
PT nutritional value
PS Claim 14; Fig 1A-H; 37pp; English.
CC This is the nucleotide sequence of a clone of the fluory2 (fl2)
CC gene of maize. It codes for a 24-kDa alpha-zein protein (see
CC W23977) that includes a 21-amino acid signal peptide (see W23976)
CC which targets the alpha-zein to the lumen of the rough endoplasmic
CC reticulum. A claimed cereal plant contains a transgene comprising
CC a first polynucleotide that encodes the fl2 signal peptide and a
CC second polynucleotide that encodes an agronomically high-value
CC protein. Also new are seeds produced by the plants, the transgene
CC itself, and a transgene that also includes the fl2 promoter. The
CC second polynucleotide preferably encodes a protein that has a high
CC content of Met, Lys, Trp and/or Thr so that feeds from transformed
CC maize, wheat, rice, barley, millet or sorghum will have increased
CC contents of these essential amino acids in their seeds (all
CC claimed).
SQ Sequence 10529 BP; 2845 A; 2484 C; 2262 G; 2938 T;

Query Match 46.2%; Score 67; DB 1; Length 10529;
Best Local Similarity 100.0%; Pred. No. 3.3e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGGTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
DB 726 GTGGATCCCCGGGTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 667
QY 106 GGGGGCC 112
DB 666 GGGGGCC 660

RESULT 12
X24730
ID X24730 standard; DNA: 12814 BP.
AC X24730;

21-JUN-1999 (first entry)
DE Swedish-FAD APP targeting vector pMTI-2398.
KW Amyloid precursor protein; APP; human; gene targeting;
KW homologous recombination; transgenic mouse; transgenic animal;
KW animal model; Alzheimer's disease; vector; pMTI-2398;
KW Swedish-FAD APP; mutation; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT mat_peptide 1932..2276
FT /tag= a
FT /standard_name= Swedish-FAD APP
FT /note= "encodes W97997"
FT mat_peptide 5369..6160
FT /tag= b
FT /note= "encodes W97998"
FT W09909150-A1.
PN 25-FEB-1999.
PD 18-AUG-1997; U14507.
PF 18-AUG-1997; WO-014507.
PR 18-AUG-1997; BAYER CORP.
PA (FARB ) BAYER CORP.
PI Wirak DO;
DR WPI: 99-181029/15.
PR P-PSDB: W97997, W97998.
DE Modification of target nucleic acids - by homologous recombination,
PT used particularly for introducing a humanised amyloid precursor
PT protein gene into rodents for producing models of Alzheimer's
PT disease
PT Example: Page 56-103; 209pp; English.
CC This is the nucleotide sequence of Swedish-FAD APP targeting
CC vector pMTI-2398. The invention provides a novel gene targeting
CC strategy that facilitates the introduction of one or more specific
CC mutations into any gene in a single double reciprocal homologous
CC recombination step. The method has been used particularly for
CC introducing a humanised amyloid precursor protein (APP) gene into
CC rodents for producing animal models of Alzheimer's disease (AD).
CC 4 independent lines of transgenic mice (lines ES5007, ES5103,
CC ES5401 and ES5403) have been created using the gene targeting
CC technique applied to embryonic stem cells. In each line, the mouse
CC APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC amino acid residues 666-770 of APP770 were encoded by human CDNA
CC sequences instead of mouse genomic exons (exons 16-18). Within
CC these residues, only 3 amino acid differences exist between the
CC mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
CC containing a humanised beta-amyloid domain. In line ES5007,
CC created using vector pMTI-2398, the Swedish mutation, i.e.
CC KM(670,681)NL, was also introduced. The targeted Swedish-FAD
CC m/hAPP gene expressed m/hAPP protein at levels approaching those
CC observed for mouse APP in brain.
SQ Sequence 12814 BP; 1865 A; 1850 C; 1962 G; 1939 T;

Query Match 46.2%; Score 67; DB 1; Length 12814;
Best Local Similarity 100.0%; Pred. No. 3.5e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGGTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
DB 9888 GTGGATCCCCGGGTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 9947
QY 106 GGGGGCC 112
DB 9948 GGGGGCC 9934

RESULT 13
X24731
ID X24731 standard; DNA: 15692 BP.
AC X24731;
DE 21-JUN-1999 (first entry)
DE London-FAD APP targeting vector pMTI-2453.
KW Amyloid precursor protein; APP; human; gene targeting;

```


KW homologous recombination; transgenic mouse; transgenic animal;
KW animal model; Alzheimer's disease; vector; pMTI-2455;
OS Swedish-FAD APP713; mutation; ss.
OS Homo sapiens.
OS Synthetic.

FT Key Location/Qualifiers
FT mat_peptide 4807..4983
FT /tag= a
FT /note= "encodes W98001"
FT mutation replace(4835,"")
FT /tag= b
FT /standard_name= "Swedish-FAD"
FT mutation replace(4981,"")
FT /tag= c
FT /standard_name= APP713stop
FT mat_peptide 8232..9032
FT /tag= d
FT /note= "encodes W97998"

PN W09090150-A1.

PD 25-FEB-1999.

PE 18-AUG-1997; U14507.

PR 18-AUG-1997; WO-U14507.

PA (FARB) BAYER CORP.

PI Wirak DO.

DR WPI; 99-181029/15.

DR P-PSDB; W97998, W98001.

PT Modification of target nucleic acids - by homologous recombination,
PT used particularly for introducing a humanised amyloid precursor
PT protein gene into rodents for producing models of Alzheimer's
PT disease

PS Example: Page 124-133; 209pp; English.

CC This is the nucleotide sequence of Swedish-FAD APP713 targeting
CC vector pMTI-2455. The invention provides a novel gene targeting
CC strategy that facilitates the introduction of one or more specific
CC mutations into any gene in a single double reciprocal homologous
CC recombination step. The method has been used particularly for
CC introducing a humanised amyloid precursor protein (APP) gene into
CC rodents for producing animal models of Alzheimer's disease (AD).
CC 4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC ES5401 and ES5403) have been created using the gene targeting
CC technique applied to embryonic stem cells. In each line, the mouse
CC APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC amino acid residues 666-770 of APP770 were encoded by human cDNA
CC sequences instead of mouse genomic exons (exons 16-18). Within
CC these residues, only 3 amino acid differences exist between the
CC mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
CC containing a humanised beta-amyloid domain. In line ES5215,
CC created using vector pMTI-2455, the Swedish mutation. i.e.
CC KM(670,681)NL, was also introduced, and also a stop codon (T to
CC stop at position 714).

SQ Sequence 15701 BP; 2470 A; 2675 C; 2845 G; 2497 T;

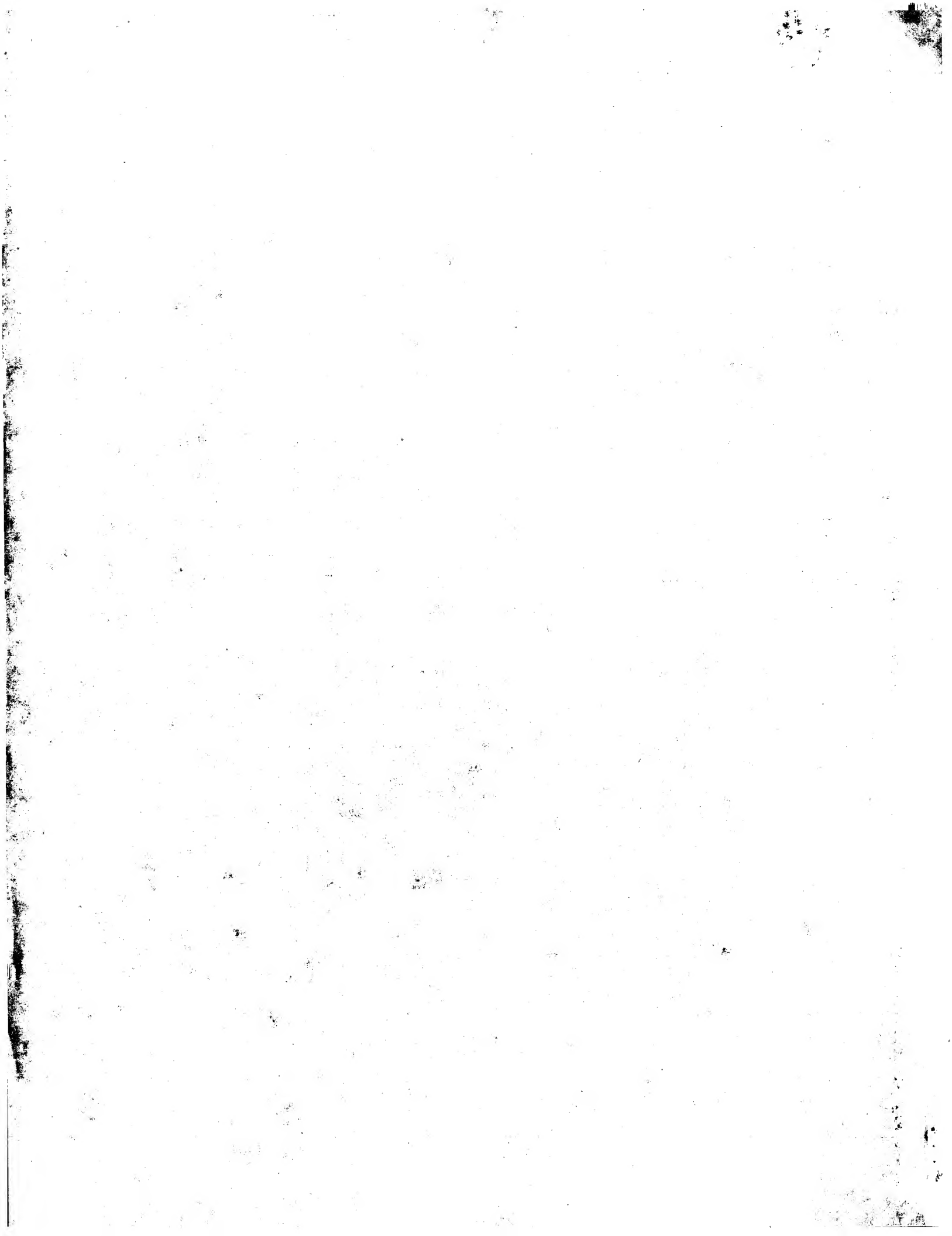
Query Match 46.2%; Score 67; DB 1; Length 15701;
Best Local Similarity 100.0%; Pred. No. 3.8e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTATTCGATACCGTCGACCTCGAGG 105
|||||
DB 12775 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTATTCGATACCGTCGACCTCGAGG 12834
|||||

QY 106 GGGGGCC 112
|||||

DB 12835 GGGGGCC 12841

Search completed: May 29, 2000, 21:58:21
Job time: 38169 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 22:08:54 ; Search time 621.83 Seconds
(without alignments)
30.310 Million cell updates/sec

Title: US-08-935-377-6
Perfect score: 145
Sequence: 1 GCGCAAAATGAAAAACTA.....TTGTTTTTGTGGCCCGCGGCC 145

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 450926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgnl_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgnl_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgnl_6/ptodata/1/ina/5C_COMB.seq.*
4: /cgnl_6/ptodata/1/ina/5D_COMB.seq.*
5: /cgnl_6/ptodata/1/ina/6_COMB.seq.*
6: /cgnl_6/ptodata/1/ina/PCFUS_COMB.seq.*
7: /cgnl_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	68.2	47.0	793	US-08-967-101-121	Sequence 121, App
C 2	68.2	47.0	793	US-08-967-101-121	Sequence 121, App
C 3	67	46.2	3792	US-08-992-334-1	Sequence 1, Appli
C 4	67	46.2	3792	US-08-302-752-1	Sequence 1, Appli
C 5	67	46.2	4016	US-08-410-540-3	Sequence 3, Appli
C 6	67	46.2	5234	US-08-992-334-2	Sequence 2, Appli
C 7	67	46.2	5234	US-08-302-752-2	Sequence 2, Appli
C 8	67	46.2	6722	US-08-992-334-3	Sequence 3, Appli
C 9	67	46.2	6722	US-08-302-752-3	Sequence 3, Appli
C 10	65.4	45.1	9318	US-08-793-610-6	Sequence 6, Appli
C 11	62.6	43.2	1023	US-08-198-446B-1	Sequence 1, Appli
C 12	62.6	43.2	1023	US-08-870-693-1	Sequence 1, Appli
C 13	62.6	43.2	2150	US-08-198-446B-10	Sequence 10, Appli
C 14	62.6	43.2	2150	US-08-870-693-10	Sequence 10, Appli
C 15	62.4	43.0	685	US-08-463-113-36	Sequence 56, Appli
C 16	62.4	43.0	685	US-08-465-388-56	Sequence 56, Appli
C 17	59	40.7	4164	US-08-204-675-1	Sequence 1, Appli
C 18	59	40.7	4164	US-08-660-754-1	Sequence 1, Appli
C 19	59	40.7	4164	US-08-798-364-1	Sequence 1, Appli
C 20	59	40.7	4164	PCT-US95-02520-1	Sequence 1, Appli
C 21	58	40.0	5178	US-08-474-169-2	Sequence 2, Appli
C 22	56	38.6	88	US-08-144-602B-15	Sequence 15, Appli
C 23	55	37.9	834	US-08-967-101-113	Sequence 113, App
C 24	55	37.9	834	US-08-592-541-113	Sequence 113, App
C 25	55	37.9	6505	US-08-793-610-5	Sequence 5, Appli
C 26	54	37.2	5534	US-08-452-267-3	Sequence 3, Appli

27	50.6	34.9	3341	3	US-08-868-577-18	Sequence 18, Appli
C 28	50	34.5	780	3	US-08-967-101-116	Sequence 116, App
C 29	50	34.5	780	4	US-08-592-541-116	Sequence 116, App
C 30	50	34.5	2308	1	US-08-325-071-62	Sequence 62, Appli
C 31	49.8	34.3	1200	1	US-08-011-398B-3	Sequence 3, Appli
C 32	49.8	34.3	1200	2	US-08-454-051-3	Sequence 3, Appli
C 33	49.8	34.3	1200	3	US-08-452-498-3	Sequence 3, Appli
C 34	49.8	34.3	6244	1	US-08-076-726-15	Sequence 15, Appli
C 35	49.8	34.3	6244	1	US-08-260-452-8	Sequence 8, Appli
C 36	49.8	34.3	6244	3	US-08-481-970-8	Sequence 8, Appli
C 37	49.8	34.3	6244	4	US-08-897-719-8	Sequence 8, Appli
C 38	48	33.1	1612	1	US-08-343-733A-1	Sequence 1, Appli
C 39	48	33.1	2764	4	US-08-455-971B-1	Sequence 1, Appli
C 40	47.2	32.6	2185	4	US-08-467-948A-3	Sequence 3, Appli
C 41	47	32.4	1997	4	US-08-667-809B-3	Sequence 3, Appli
C 42	47	32.4	2126	3	US-08-789-354-1	Sequence 1, Appli
C 43	47	32.4	2605	3	US-08-680-395-4	Sequence 4, Appli
C 44	46.6	32.1	1249	2	US-08-463-115-35	Sequence 35, Appli
C 45	46.6	32.1	1249	2	US-08-455-388-35	Sequence 35, Appli

ALIGNMENTS

RESULT 1
US-08-967-101-121/c
; Sequence 121, Application US/08967101
; Patent No. 5840540
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,101
; FILING DATE: 10-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA-(genomic)
US-08-967-101-121

Query Match 47.0%; Score 68.2; DB 3; Length 793;
Best Local Similarity 85.4%; Pred. No. 2.7e-14;
Matches 76; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 42 CGCGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 101
Db 122 CGGGGGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 63
QY 102 GAGGGGGGCTTAACCTAATTAATTTGTT 130
Db 62 GAGGGGGGCGCGGTACCCAGCTTTTGT 34

RESULT 2

US-08-592-541-121/c
; Sequence 121, Application US/08592541
; Patent No. 5986054
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMWENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-592-541-121

Query Match 47.0%; Score 68.2; DB 4; Length 793;
Best Local Similarity 85.4%; Pred. No. 2.7e-14;
Matches 76; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 42 CGCGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 101
Db 122 CGGGGGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 63
QY 102 GAGGGGGGCGCGGTAACTAATTAATTTGTT 130
Db 62 GAGGGGGGCGCGGTACCCAGCTTTTGT 34

RESULT 3

US-08-992-334-1/c
; Sequence 1, Application US/08992334
; Patent No. 5919678
; GENERAL INFORMATION:
; APPLICANT: Gruss, Alexandra
; APPLICANT: Maguin, Emmanuelle
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE
; TITLE OF INVENTION: PLASMID

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christie Parker & Hale, LLP
STREET: 350 West Colorado Boulevard, Suite 500
CITY: Pasadena
STATE: California
COUNTRY: United States
ZIP: 91105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/992,334
FILING DATE: 17-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302,752
FILING DATE: 24-DEC-1994
APPLICATION NUMBER: PCT/FR93/00248
FILING DATE: 12-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/03034
FILING DATE: 13-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Prout, D. Bruce
REGISTRATION NUMBER: 20958
REFERENCE/DOCKET NUMBER: C93:31779
TELECOMMUNICATION INFORMATION:
TELEPHONE: (626) 795-9900
TELEFAX: (626) 577-8800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3792 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: YES
IMMEDIATE SOURCE:
CLONE: pg+host4
US-08-992-334-1

Query Match 46.2%; Score 67; DB 4; Length 3792;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGATATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
Db 3541 GTGATATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 3482
QY 106 GGGGGGCC 112
Db 3481 GGGGGGCC 3475

RESULT 4

US-08-302-752-1/c
; Sequence 1, Application US/08302752
; Patent No. 6025190
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: THERMOSENSIBLE PLASMID
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,752
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9203034
; FILING DATE: 13-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR/93/00248
; FILING DATE: 12-MAR-1993
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3792 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-302-752-1

Query Match 46.2%; Score 67; DB 5; Length 3792;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
DB 3541 GTGGATCCCCGGCTGCAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 3482
QY 106 GGGGGCC 112
DB 3481 GGGGGCC 3475

RESULT 5
US-08-410-540-3
; Sequence 3, Application US/08410540
; Patent No. 5807678
; GENERAL INFORMATION:
; APPLICANT: Miller, Walter L.
; APPLICANT: Lin, Dong
; APPLICANT: Strauss III, Jerome F.
; TITLE OF INVENTION: IDENTIFICATION OF GENE MUTATIONS
; TITLE OF INVENTION: ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Cooley Godward Castro Huddleson & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94306-2155
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/410,540
; FILING DATE: 23-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Neeley, Richard L.
; REGISTRATION NUMBER: 30,092
; REFERENCE/DOCKET NUMBER: UCAL-238/0005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 853 5070
; TELEFAX: 415 857 0663
; TELEX: 380816COOLEIPA
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1098..1283
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1620..1733
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2047..2174
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2267..2425
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2567..2751
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2828..2921
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3031..3765
; FEATURE:
; NAME/KEY: -
; LOCATION: 1433..1434
; OTHER INFORMATION: /note= "interruption of sequence data"
; FEATURE:
; NAME/KEY: -
; LOCATION: 2208..2209
; OTHER INFORMATION: /note= "interruption of sequence data"
; FEATURE:
; NAME/KEY: -
; LOCATION: 2781..2782
; OTHER INFORMATION: /note= "interruption of sequence data"
; FEATURE:
; NAME/KEY: -
; LOCATION: 2956..2957
; OTHER INFORMATION: /note= "interruption of sequence data"
; US-08-410-540-3

Query Match 46.2%; Score 67; DB 2; Length 4016;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
DB 3939 GTGGATCCCCGGCTGCAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 3998
QY 106 GGGGGCC 112
DB 3999 GGGGGCC 4005

RESULT 6
US-08-992-334-2/c
; Sequence 2, Application US/08992334
; Patent No. 5919678
; GENERAL INFORMATION:
; APPLICANT: Gruss, Alexandra
; APPLICANT: Maguin, Emmanuelle
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE
; TITLE OF INVENTION: PLASMID
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christie Parker & Hale, LLP
; STREET: 350 West Colorado Boulevard, Suite 500
; CITY: Pasadena
; STATE: California
; COUNTRY: United States
```

```
;
;
; ZIP: 91105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992,334
; FILING DATE: 17-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,752
; FILING DATE: 24-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00248
; FILING DATE: 12-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/03034
; FILING DATE: 13-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Prout, D. Bruce
; REGISTRATION NUMBER: 20958
; REFERENCE/DOCKET NUMBER: C93:31779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626) 795-9900
; TELEFAX: (626) 577-8800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5234 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-992-334-2

Query Match          46.2%; Score 67; DB 4; Length 5234;
Best Local Similarity 100.0%; Pred. No. 1.2e-13;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
DB 4983 GTGGATCCCCGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4924

QY 106 GGGGGCC 112
DB 4923 GGGGGCC 4917

RESULT 7
US-08-302-752-2/c
; Sequence 2, Application US/08302752
; Patent No. 6025190
; GENERAL INFORMATION:
; APPLICANT: THERMOSENSIBLE PLASMID
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,752
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9203034
; FILING DATE: 13-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR/93/00248
; FILING DATE: 12-MAR-1993
; INFORMATION FOR SEQ ID NO: 2:
```

```
;
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5234 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-302-752-2

Query Match          46.2%; Score 67; DB 5; Length 5234;
Best Local Similarity 100.0%; Pred. No. 1.2e-13;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
DB 4983 GTGGATCCCCGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4924

QY 106 GGGGGCC 112
DB 4923 GGGGGCC 4917

RESULT 8
US-08-992-334-3/c
; Sequence 3, Application US/08992334
; Patent No. 5919678
; GENERAL INFORMATION:
; APPLICANT: Gruss, Alexandra
; APPLICANT: Maguin, Emmanuelle
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE
; TITLE OF INVENTION: PLASMID
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christie Parker & Hale, LLP
; STREET: 350 West Colorado Boulevard, Suite 500
; CITY: Pasadena
; STATE: California
; COUNTRY: United States
; ZIP: 91105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992,334
; FILING DATE: 17-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,752
; FILING DATE: 24-DEC-1994
; PRIOR APPLICATION DATA: PCT/FR93/00248
; APPLICATION NUMBER:
; FILING DATE: 12-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/03034
; FILING DATE: 13-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Prout, D. Bruce
; REGISTRATION NUMBER: 20958
; REFERENCE/DOCKET NUMBER: C93:31779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626) 795-9900
; TELEFAX: (626) 577-8800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-992-334-3
```



```
;
;
; Query Match 46.2%; Score 67; DB 4; Length 6722;
; Best Local Similarity 100.0%; Pred. No. 1.3e-13;
; Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
Db 6471 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 6412

Qy 106 GGGGGCC 112
Db 6411 GGGGGCC 6405

RESULT 9
US-08-302-752-3/c
; Sequence 3, Application US/08302752
; Patent No. 6025190
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: THERMOSENSIBLE PLASMID
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,752
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9203034
; FILING DATE: 13-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR/93/00248
; FILING DATE: 12-MAR-1993
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-302-752-3

Query Match 46.2%; Score 67; DB 5; Length 6722;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
Db 6471 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 6412

Qy 106 GGGGGCC 112
Db 6411 GGGGGCC 6405

RESULT 10
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BADM, Christopher
; APPLICANT: STOCKING-HARRIS, Carol
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
;
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-08-793-610-6

Query Match 45.1%; Score 65.4; DB 3; Length 9318;
Best Local Similarity 98.5%; Pred. No. 4.9e-13;
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
Db 5634 GGGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 5713

Qy 106 GGGGGCC 112
Db 5714 GGGGGCC 5720

RESULT 11
US-08-198-446B-1/c
; Sequence 1, Application US/08198446B
; Patent No. 5674996
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Plon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/198,446B
; APPLICATION NUMBER: US/08/198,446B
; FILING DATE: 18-FEB-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheiness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: FHCRI17537
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100
; TELEFAX: 206-224-0779
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1023 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; DESCRIPTION: yeast RAD17 CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Saccharomyces cerevisiae
;
; US-08-198-446B-1
;
; Query Match 43.2%; Score 62.6; DB 1; Length 1023;
; Best Local Similarity 73.4%; Pred. No. 2.2e-12;
; Matches 80; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
;
; QY 4 CAAAATTCGAAATCAATCATCTATTATTCGACGGCGCGCGGTGGATCCCGCGGGCTGC 63
; DB 204 CGAAATGATAAATATGGGAATTTGTCATTGACAAATGCGATTTAGGGGCTGC 145
;
; QY 64 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCC 112
; DB 144 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCC 96
;
; RESULT 12
; US-08-870-693-1/c
; Sequence 1, Application US/08870693
; Patent No. 5866338
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Plon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/870,693
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/198,446
; FILING DATE: February 18, 1994
; APPLICATION NUMBER: PCT/US93/04458
; FILING DATE: May 12, 1993
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/198,446B
; APPLICATION NUMBER: US/08/198,446B
; FILING DATE: 18-FEB-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheiness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: FHCRI10798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100
; TELEFAX: 206-224-0779
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1023 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; DESCRIPTION: yeast RAD17 CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Saccharomyces cerevisiae
;
; US-08-870-693-1
;
; Query Match 43.2%; Score 62.6; DB 3; Length 1023;
; Best Local Similarity 73.4%; Pred. No. 2.2e-12;
; Matches 80; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
;
; QY 4 CAAAATTCGAAATCAATCATCTATTATTCGACGGCGCGCGGTGGATCCCGCGGGCTGC 63
; DB 204 CGAAATGATAAATATGGGAATTTGTCATTGACAAATGCGATTTAGGGGCTGC 145
;
; QY 64 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCC 112
; DB 144 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCC 96
;
; RESULT 13
; US-08-198-446B-10/c
; Sequence 10, Application US/08198446B
; Patent No. 5674996
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Plon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/198,446B
; FILING DATE: 18-FEB-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheiness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: FHCRI17537
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100
; TELEFAX: 206-224-0779
;

```

TELEFAX: 206-224-0779
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2150 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: Yeast RAD17 cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharomyces cerevisiae
FEATURE:
NAME/KEY: CDS
LOCATION: 741..1947
US-08-198-446B-10

Query Match 43.2%; Score 62.6; DB 1; Length 2150;
Best Local Similarity 73.4%; Pred. No. 2.8e-12;
Matches 80; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Qy 4 CAAAAATTGAAACTAGATCTATTATTGACGCGCGCGCGTGGATCCCCCGGGCTGC 63
Db 204 CGAAATGATAAAATTGGAAATTGATCCATTGACAAATGCAATTTAGGGGCTGC 145
Qy 64 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCC 112
Db 144 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCC 96

RESULT 14
US-08-870-693-10/c
Sequence 10, Application US/08870693
Patent No. 5866338
GENERAL INFORMATION:
APPLICANT: Hartwell, Leland H.
APPLICANT: Weinert, Ted A.
APPLICANT: Plon, Sharon E.
APPLICANT: Groudine, Mark T.
TITLE OF INVENTION: Cell Cycle Checkpoint Genes
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
STREET: 1420 Fifth Ave., Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/870,693
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/198,446
FILING DATE: February 18, 1994
APPLICATION NUMBER: PCT/US93/04458
FILING DATE: May 12, 1993
APPLICATION NUMBER: US 07/884,426
FILING DATE: May 14, 1992
APPLICATION NUMBER: US 07/882,051
FILING DATE: May 12, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHCRI10798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100

TELEFAX: 206-224-0779
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2150 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: Yeast RAD17 cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharomyces cerevisiae
FEATURE:
NAME/KEY: CDS
LOCATION: 741..1947
US-08-870-693-10

Query Match 43.2%; Score 62.6; DB 3; Length 2150;
Best Local Similarity 73.4%; Pred. No. 2.8e-12;
Matches 80; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Qy 4 CAAAAATTGAAACTAGATCTATTATTGACGCGCGCGCGTGGATCCCCCGGGCTGC 63
Db 204 CGAAATGATAAAATTGGAAATTGATCCATTGAAACAAATGCAATTTAGGGGCTGC 145
Qy 64 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCC 112
Db 144 AGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCC 96

RESULT 15
US-08-463-115-56
Sequence 56, Application US/08463115
Patent No. 5703221
GENERAL INFORMATION:
APPLICANT: WILLIAM JOHN MARTIN
TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
AND RELATED VACCINES
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,115
FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/157,811
FILING DATE: No. 5703221ember 23, 1993
APPLICATION NUMBER: 07/887,502
FILING DATE: May 22, 1992
APPLICATION NUMBER: 07/704,814
FILING DATE: May 23, 1991
APPLICATION NUMBER: 07/763,039
FILING DATE: September 20, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327

```
; REFERENCE/DOCKET NUMBER: 213/301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 685 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION:
; US-08-463-115-56

Query Match      43.0%; Score 62.4; DB 2; Length 685;
Best Local Similarity 94.0%; Pred. No. 2.3e-12;
Matches 63; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
Db 150 GTGGATCCCCGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 209

QY 106 GGGGGCC 112
Db 210 GGGGGNC 216
```

Search completed: May 29, 2000, 22:09:00
Job time: 38741 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2000, 09:48:27 ; Search time 2276.24 Seconds
(without alignments)
284.222 Million cell updates/sec

Title: US-08-935-377-6

Perfect score: 145
Sequence: 1 GGCCAAAATTGAAACTA.....TTGTTTTGTGGCCCGGCC 145

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5142629 seqs, 2230885800 residues

Total number of hits satisfying chosen parameters: 10285240

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA.*

1: /cgnl_6/ptodata/1/pna/US092B_COMB.seq.*
2: /cgnl_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgnl_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgnl_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgnl_6/ptodata/1/pna/US081A_COMB.seq.*
6: /cgnl_6/ptodata/1/pna/US081B_COMB.seq.*
7: /cgnl_6/ptodata/1/pna/US081C_COMB.seq.*
8: /cgnl_6/ptodata/1/pna/US081D_COMB.seq.*
9: /cgnl_6/ptodata/1/pna/US082A_COMB.seq.*
10: /cgnl_6/ptodata/1/pna/US082B_COMB.seq.*
11: /cgnl_6/ptodata/1/pna/US082C_COMB.seq.*
12: /cgnl_6/ptodata/1/pna/US083A_COMB.seq.*
13: /cgnl_6/ptodata/1/pna/US083B_COMB.seq.*
14: /cgnl_6/ptodata/1/pna/US084A_COMB.seq.*
15: /cgnl_6/ptodata/1/pna/US084B_COMB.seq.*
16: /cgnl_6/ptodata/1/pna/US084C_COMB.seq.*
17: /cgnl_6/ptodata/1/pna/US085_COMB.seq.*
18: /cgnl_6/ptodata/1/pna/US086_COMB.seq.*
19: /cgnl_6/ptodata/1/pna/US087A_COMB.seq.*
20: /cgnl_6/ptodata/1/pna/US087B_COMB.seq.*
21: /cgnl_6/ptodata/1/pna/US087C_COMB.seq.*
22: /cgnl_6/ptodata/1/pna/US088A_COMB.seq.*
23: /cgnl_6/ptodata/1/pna/US088B_COMB.seq.*
24: /cgnl_6/ptodata/1/pna/US088C_COMB.seq.*
25: /cgnl_6/ptodata/1/pna/US088D_COMB.seq.*
26: /cgnl_6/ptodata/1/pna/US089A_COMB.seq.*
27: /cgnl_6/ptodata/1/pna/US089B_COMB.seq.*
28: /cgnl_6/ptodata/1/pna/US089C_COMB.seq.*
29: /cgnl_6/ptodata/1/pna/US089D_COMB.seq.*
30: /cgnl_6/ptodata/1/pna/US089E_COMB.seq.*
31: /cgnl_6/ptodata/1/pna/US089F_COMB.seq.*
32: /cgnl_6/ptodata/1/pna/US089G_COMB.seq.*
33: /cgnl_6/ptodata/1/pna/US089H_COMB.seq.*
34: /cgnl_6/ptodata/1/pna/US089I_COMB.seq.*
35: /cgnl_6/ptodata/1/pna/US090A_COMB.seq.*
36: /cgnl_6/ptodata/1/pna/US090B_COMB.seq.*
37: /cgnl_6/ptodata/1/pna/US090C_COMB.seq.*
38: /cgnl_6/ptodata/1/pna/US090D_COMB.seq.*
39: /cgnl_6/ptodata/1/pna/US090E_COMB.seq.*
40: /cgnl_6/ptodata/1/pna/US091A_COMB.seq.*
41: /cgnl_6/ptodata/1/pna/US091B_COMB.seq.*
42: /cgnl_6/ptodata/1/pna/US091C_COMB.seq.*
43: /cgnl_6/ptodata/1/pna/US092A_COMB.seq.*

44: /cgnl_6/ptodata/1/pna/US092B_COMB.seq.*
45: /cgnl_6/ptodata/1/pna/US092C_COMB.seq.*
46: /cgnl_6/ptodata/1/pna/US092D_COMB.seq.*
47: /cgnl_6/ptodata/1/pna/US092E_COMB.seq.*
48: /cgnl_6/ptodata/1/pna/US093A_COMB.seq.*
49: /cgnl_6/ptodata/1/pna/US093B_COMB.seq.*
50: /cgnl_6/ptodata/1/pna/US093C_COMB.seq.*
51: /cgnl_6/ptodata/1/pna/US093D_COMB.seq.*
52: /cgnl_6/ptodata/1/pna/US093E_COMB.seq.*
53: /cgnl_6/ptodata/1/pna/US093F_COMB.seq.*
54: /cgnl_6/ptodata/1/pna/US094A_COMB.seq.*
55: /cgnl_6/ptodata/1/pna/US094B_COMB.seq.*
56: /cgnl_6/ptodata/1/pna/US094C_COMB.seq.*
57: /cgnl_6/ptodata/1/pna/US095_COMB.seq.*
58: /cgnl_6/ptodata/1/pna/US6000_COMB.seq.*
59: /cgnl_6/ptodata/1/pna/US6001A_COMB.seq.*
60: /cgnl_6/ptodata/1/pna/US6001B_COMB.seq.*
61: /cgnl_6/ptodata/1/pna/US6002A_COMB.seq.*
62: /cgnl_6/ptodata/1/pna/US6002B_COMB.seq.*
63: /cgnl_6/ptodata/1/pna/US6002C_COMB.seq.*
64: /cgnl_6/ptodata/1/pna/US6003A_COMB.seq.*
65: /cgnl_6/ptodata/1/pna/US6003B_COMB.seq.*
66: /cgnl_6/ptodata/1/pna/US6004A_COMB.seq.*
67: /cgnl_6/ptodata/1/pna/US6004B_COMB.seq.*
68: /cgnl_6/ptodata/1/pna/US6004C_COMB.seq.*
69: /cgnl_6/ptodata/1/pna/US6005_COMB.seq.*
70: /cgnl_6/ptodata/1/pna/US6006_COMB.seq.*
71: /cgnl_6/ptodata/1/pna/US6007_COMB.seq.*
72: /cgnl_6/ptodata/1/pna/US6008A_COMB.seq.*
73: /cgnl_6/ptodata/1/pna/US6008B_COMB.seq.*
74: /cgnl_6/ptodata/1/pna/US6008C_COMB.seq.*
75: /cgnl_6/ptodata/1/pna/US6009A_COMB.seq.*
76: /cgnl_6/ptodata/1/pna/US6009B_COMB.seq.*
77: /cgnl_6/ptodata/1/pna/US6010A_COMB.seq.*
78: /cgnl_6/ptodata/1/pna/US6010B_COMB.seq.*
79: /cgnl_6/ptodata/1/pna/US6011A_COMB.seq.*
80: /cgnl_6/ptodata/1/pna/US6011B_COMB.seq.*
81: /cgnl_6/ptodata/1/pna/US6012A_COMB.seq.*
82: /cgnl_6/ptodata/1/pna/US6012B_COMB.seq.*
83: /cgnl_6/ptodata/1/pna/US6013A_COMB.seq.*
84: /cgnl_6/ptodata/1/pna/US6013B_COMB.seq.*
85: /cgnl_6/ptodata/1/pna/US6014A_COMB.seq.*
86: /cgnl_6/ptodata/1/pna/US6014B_COMB.seq.*
87: /cgnl_6/ptodata/1/pna/US6014C_COMB.seq.*
88: /cgnl_6/ptodata/1/pna/US6015_COMB.seq.*
89: /cgnl_6/ptodata/1/pna/US6016A_COMB.seq.*
90: /cgnl_6/ptodata/1/pna/US6016B_COMB.seq.*
91: /cgnl_6/ptodata/1/pna/US6016C_COMB.seq.*
92: /cgnl_6/ptodata/1/pna/US6017A_COMB.seq.*
93: /cgnl_6/ptodata/1/pna/US6017B_COMB.seq.*
94: /cgnl_6/ptodata/1/pna/US6017C_COMB.seq.*
95: /cgnl_6/ptodata/1/pna/US6018A_COMB.seq.*
96: /cgnl_6/ptodata/1/pna/US6018B_COMB.seq.*
97: /cgnl_6/ptodata/1/pna/US6018C_COMB.seq.*
98: /cgnl_6/ptodata/1/pna/US6019_COMB.seq.*
99: /cgnl_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
100: /cgnl_6/ptodata/1/pna/US06_NEW_COMB.seq.*
101: /cgnl_6/ptodata/1/pna/US07_NEW_COMB.seq.*
102: /cgnl_6/ptodata/1/pna/US08_NEW_COMB.seq.*
103: /cgnl_6/ptodata/1/pna/US09_NEW_COMB.seq.*
104: /cgnl_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	145	100.0	145	28	US-08-935-377-6
2	132	91.0	148	28	US-08-935-377-7

3	131	90.3	149	28	US-08-935-377-8	Sequence 8, Appl1
4	130	89.7	150	28	US-08-935-377-9	Sequence 9, Appl1
5	69	47.6	224	13	US-08-377-383-4	Sequence 4, Appl1
6	69	47.3	224	18	US-08-693-573-4	Sequence 4, Appl1
7	68.6	47.3	858	40	US-09-123-912-91	Sequence 91, Appl1
8	68.6	47.3	858	43	US-09-221-107-91	Sequence 91, Appl1
9	68.6	47.3	858	46	US-09-285-479-91	Sequence 91, Appl1
10	68.2	47.0	793	15	US-08-431-048A-121	Sequence 121, App
11	68.2	47.0	793	16	US-08-496-841-121	Sequence 121, App
12	68.2	47.0	793	17	US-08-509-359-121	Sequence 121, App
13	68.2	47.0	793	17	US-08-509-359B-121	Sequence 121, App
14	68.2	47.0	793	40	US-09-124-523-121	Sequence 121, App
15	68.2	47.0	793	40	US-09-124-698-121	Sequence 121, App
16	68.2	47.0	793	40	US-09-127-480-121	Sequence 121, App
17	67	46.2	147	38	US-09-072-433-31	Sequence 31, Appl
18	67	46.2	157	63	US-60-029-960-595	Sequence 595, App
19	67	46.2	245	54	US-09-411-999-19130	Sequence 19130, A
20	67	46.2	381	49	US-09-332-782-8570	Sequence 8570, Ap
21	67	46.2	381	57	US-09-515-694-8570	Sequence 8570, Ap
22	67	46.2	393	54	US-09-411-999-13627	Sequence 13627, A
23	67	46.2	405	49	US-09-332-782-8823	Sequence 8523, Ap
24	67	46.2	405	57	US-09-515-694-8823	Sequence 8523, Ap
25	67	46.2	420	49	US-09-332-782-8595	Sequence 8595, Ap
26	67	46.2	420	57	US-09-515-694-8595	Sequence 8595, Ap
27	67	46.2	422	49	US-09-332-782-9309	Sequence 9309, Ap
28	67	46.2	422	57	US-09-515-694-9309	Sequence 9309, Ap
29	67	46.2	423	54	US-09-411-999-34199	Sequence 34199, A
30	67	46.2	433	42	US-09-179-473-2334	Sequence 2334, Ap
31	67	46.2	433	49	US-09-328-351-2334	Sequence 2334, Ap
32	67	46.2	442	42	US-09-179-473-2495	Sequence 2495, Ap
33	67	46.2	442	49	US-09-328-351-2495	Sequence 2495, Ap
34	67	46.2	448	42	US-09-179-473-2255	Sequence 2255, Ap
35	67	46.2	448	49	US-09-328-351-2255	Sequence 2255, Ap
36	67	46.2	454	42	US-09-179-473-2100	Sequence 2100, Ap
37	67	46.2	454	42	US-09-328-351-2100	Sequence 2100, Ap
38	67	46.2	454	49	US-09-328-351-2360	Sequence 2360, Ap
39	67	46.2	454	49	US-09-328-351-2360	Sequence 2360, Ap
40	67	46.2	752	32	US-08-976-259-108	Sequence 108, App
41	67	46.2	752	62	US-60-031-626-108	Sequence 108, App
42	67	46.2	827	84	US-60-126-265-996	Sequence 996, App
43	67	46.2	2958	37	US-09-040-961-6	Sequence 6, Appl1
44	67	46.2	2958	41	US-09-130-115-6	Sequence 6, Appl1
45	67	46.2	2961	15	US-08-446-935-6	Sequence 6, Appl1

Query Match 100.0% Score 145; DB 28; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.3e-38;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGCCAAATTCGAAACTAGATCTATTATTCACGCGCGCGCGTGGATCCCCCGGC	60
Db	1	GGCCAAATTCGAAACTAGATCTATTATTCACGCGCGCGCGTGGATCCCCCGGC	60
QY	61	TCGAGGAATTCGATATCAAGCTTATCGATACCGTGCAGGCGGCGGCGCTAACTAAC	120
Db	61	TCGAGGAATTCGATATCAAGCTTATCGATACCGTGCAGGCGGCGGCGCTAACTAAC	120
QY	121	TAATTTGTTTGTGGCGCGGCC	145
Db	121	TAATTTGTTTGTGGCGCGGCC	145

US-08-935-377-7

RESULT 2

US-08-935-377-7

Sequence 7, Application US/08935377

GENERAL INFORMATION:

APPLICANT: Zauderer, Maurice

TITLE OF INVENTION: T Cells Specific for Target Antigens and

TITLE OF INVENTION: Vaccines Based Thereon

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C

STREET: 1100 New York Avenue, N.W., Suite 600

CITY: Washington

STATE: D. C.

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/935,377

FILING DATE: 22-SEP-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Steffe, Eric K

REGISTRATION NUMBER: 36,688

REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 148 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-935-377-7

Db 121 TACTAAATTTGTTTTTGTGGCGCCGCC 149

RESULT 4

US-08-935-377-9

; Sequence 9, Application US/08935377

; GENERAL INFORMATION:

; APPLICANT: Zauderer, Maurice

; TITLE OF INVENTION: T Cells Specific for Target Antigens and

; TITLE OF INVENTION: Vaccines Based Thereon

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C

; STREET: 1100 New York Avenue, N.W., Suite 600

; CITY: Washington

; STATE: D. C.

; COUNTRY: USA

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/935,377

; FILING DATE: 22-SEP-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Steffe, Eric K

; REGISTRATION NUMBER: 36,688

; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 150 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-08-935-377-9

Query Match 89.7%; Score 130; DB 28; Length 150;

Best Local Similarity 96.7%; Pred. No. 1.4e-33;

Matches 145; Conservative 0; Mismatches 0; Indels 5; Gaps

Qy 1 GGGCAAAATTTGAAAACTAGATCTATTATTTCAGCGCGCCGC-----CGTGGATCCCC 55

Db 1 GGGCAAAATTTGAAAACTAGATCTATTATTTCAGCGCGCCGCATGAGTGGATCCCC 60

Qy 56 CGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCGCTAA 115

Db 61 CGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCGCTAA 120

Qy 116 CTAACATAATTTGTTTTTGTGGCGCCGGCC 145

Db 121 CTAACATAATTTGTTTTTGTGGCGCCGGCC 150

RESULT 5

US-08-377-383-4

; Sequence 4, Application US/08377383

; GENERAL INFORMATION:

; APPLICANT: HOSOI,Shigeru, FUKAMI,Tadashi, HIYOSHI,Makiko

; TITLE OF INVENTION: Method of Determining Base Sequence of

; TITLE OF INVENTION: Nucleic Acid

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cushman Darby & Cushman

; STREET: Ninth Floor, 1100 New York Avenue, N.W.

; CITY: Washington

```
STATE: DC
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 720Kb storage
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1 or ASCII editors
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/377,383
FILING DATE: 25-JAN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP6971/1994
FILING DATE: 26-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Glenn J. Perry
REGISTRATION NUMBER: 28458
TELEPHONE: (202)861-3000
TELEFAX: (202)822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid Synthetic DNA
US-08-377-383-4

Query Match 47.6%; Score 69; DB 13; Length 224;
Best Local Similarity 88.2%; Pred. No. 4.2e-13;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
DB 88 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 147
QY 106 GGGGCGCCCTAACTAATTTGTT 130
DB 148 GGGGCGCCGGTACCCAGCTTTGTT 172

RESULT 6
US-08-693-573-4
; Sequence 4, Application US/086933573
; GENERAL INFORMATION:
; APPLICANT: HOSOI, Shigeru,
; APPLICANT: FUKAMI, Tadashi,
; APPLICANT: KOJIMA, Makiko
; TITLE OF INVENTION: Method of Determining Base Sequence of
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABBY CUSHMAN
; ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: Ninth Floor, 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 720Kb storage
COMPUTER: IBM PC/XT/AT compatibles
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1 or ASCII editors
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/693,573
FILING DATE: 07-AUG-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/377,383

STATE: DC
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 720Kb storage
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1 or ASCII editors
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/377,383
FILING DATE: 25-JAN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP6971/1994
FILING DATE: 26-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Thomas G. Wiseman
REGISTRATION NUMBER: 35046
TELEPHONE: (202)861-3000
TELEFAX: (202)822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid Synthetic DNA
US-08-693-573-4

Query Match 47.6%; Score 69; DB 18; Length 224;
Best Local Similarity 88.2%; Pred. No. 4.2e-13;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
DB 88 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 147
QY 106 GGGGCGCCCTAACTAATTTGTT 130
DB 148 GGGGCGCCGGTACCCAGCTTTGTT 172

RESULT 7
US-09-123-912-91
; Sequence 91, Application US/09123912A
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C1
; CURRENT APPLICATION NUMBER: US/09/123,912A
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: 09/040,802
; PRIOR FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 91
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (570)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (591)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (635)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (664)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (667)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (683)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (711)
; OTHER INFORMATION: Where n is a, c, g or t
```


APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ & MENTLIK
STREET: 600 South Avenue West
City: Westfield
STATE: New Jersey
COUNTRY: U.S.A.

```

;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word, Version 6.0c
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/509,359B
; FILING DATE: 31-JUL-95
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Palisi, Thomas M.
; REGISTRATION NUMBER: 36629
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 654-5000
; TELEFAX: (908) 654-7866
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-509-359B-121

Query Match          47.0%   Score 68.2: DB 17:   Length 79
Best Local Similarity 85.4%   Pred. No. 1.le-12:
Matches 76: Conservative 0; Mismatches 13; Indels

Qy  42  CGCGGTGATCCCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTGC
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  122  CGGGCGGATCCCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTGC

Qy  102  GAGGGGGGGCCTAACTAACTAATTGTT 130
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  62  GAGGGGGGGCCCGGTACCCAGCTTTTGT  34

RESULT 14
us-09-124-523-121/c
Sequence 121. Application US/09124523

```

REFERENCE: CIT. GEORGE HENRIOT, JULIA H.
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: 414 High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.

```
; COMPUTER READERABLE FORM.  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/124,523
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/592,541
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Pitcher, Edmund R.
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 248-7000
;; TELEFAX: (617) 248-7100
;; INFORMATION FOR SEQ ID NO: 121:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 793 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
US-09-124-523-121

Query Match      47.0%; Score 68.2; DB 40; Length 793;
Best Local Similarity 85.4%; Pred. No. 1.le-12;
Matches 76; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 42 CGCCGTGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 101
Db 122 CGGGCGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 63

QY 102 GAGGGGGGGCCTAACTAACTAATTTGTT 130
Db 62 GAGGGGGGGCCCGGTACCCAGCTTTTGT 34

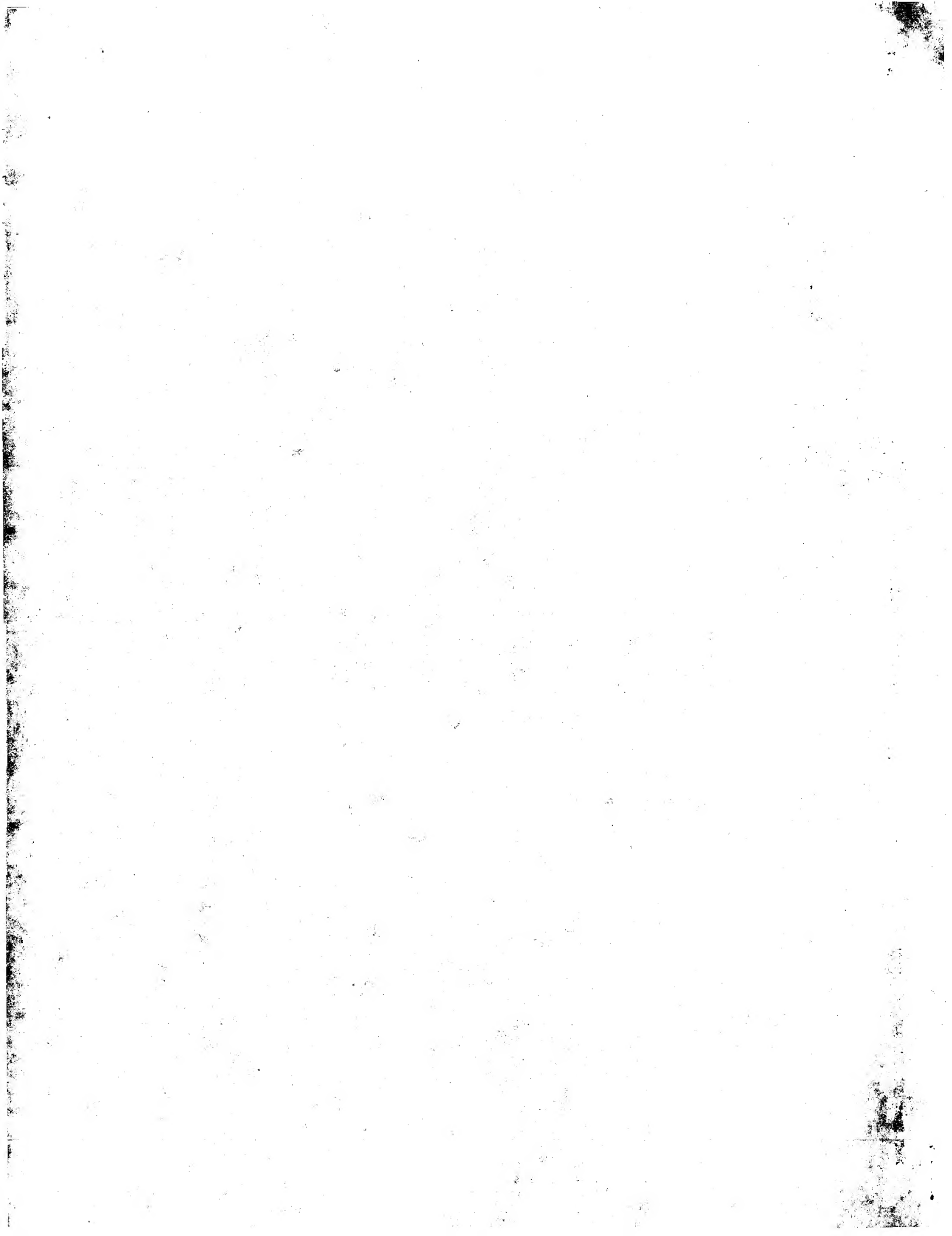
Search completed: May 30, 2000, 09:48:30
Job time: 60180 sec.
```

```
Query Match      47.0%; Score 68.2; DB 40; Length 793;
Best Local Similarity 85.4%; Pred. No. 1.le-12;
Matches 76; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 42 CGCCGTGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 101
Db 122 CGGGCGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 63

QY 102 GAGGGGGGGCCTAACTAACTAATTTGTT 130
Db 62 GAGGGGGGGCCCGGTACCCAGCTTTTGT 34
```

```
RESULT 15
US-09-124-698-121/c
Sequence 121, Application US/09124698
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,698
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 793 base pairs
```



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:13:13 ; Search time 2192.43 Seconds
(without alignments)
268.067 Million cell updates/sec

Title: US-08-935-377-6
Perfect score: 145
Sequence: 1 GGCCAAAATGAAAACTA.....TTGTTTTGTGGCCCGGCC 145

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues
Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *

44: gb_est25: *
45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: em_est20: *
50: em_est21: *
51: em_est22: *
52: gb_est30: *
53: gb_est31: *
54: gb_est32: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: gb_gss1: *
83: gb_gss2: *
84: gb_gss3: *
85: gb_gss4: *
86: em_gss1: *
87: em_gss2: *
88: em_gss3: *
89: em_gss4: *
90: gb_gss5: *
91: gb_gss6: *
92: em_gss5: *
93: em_gss6: *
94: gb_gss7: *
95: gb_gss8: *
96: gb_gss9: *
97: em_gss7: *
98: em_gss8: *
99: gb_gss11: *
100: gb_gss10: *
101: em_gss9: *
102: em_gss10: *
103: em_gss11: *
104: em_gss12: *
105: gb_gss12: *
106: gb_gss13: *
107: gb_gss14: *
108: gb_gss15: *
109: gb_gss16: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES


```
/organism="Homo sapiens"
/db_xref="GDB:7147360"
/db_xref="taxon:9606"
/clone="2281N13"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/Note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII"
HindIII
141 a 164 c 166 g 137 t

BASE COUNT      141 a 164 c 166 g 137 t
ORIGIN

Query Match      46.9%; Score 68; DB 96; Length 608;
Best Local Similarity 100.0%; Pred. No. 8.1e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGGTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
|||||
DB 82 GTGGATCCCCGGGTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 23
|||||
QY 106 GGGGGCCT 113
|||||
DB 22 GGGGGCCT 15

RESULT 3
B69688/c
LOCUS      B69688      528 bp      DNA      GSS      18-JUN-1998
DEFINITION C1978SK-A-448E10.TVC C1978SK Homo sapiens genomic clone A-448E10,
genomic survey sequence.
ACCESSION  B69688
VERSION    B69688.1 GI:2708912
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 528)
AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J.,
Shizuva,H., Simon,M. and Venter,J.C.
Shizuva,H., Simon,M. and Venter,J.C.
Use of a human BAC End Sequence Database for Sequence-Ready Map
Building
JOURNAL    Unpublished (1997)
COMMENT    Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.

FEATURES             Location/Qualifiers
     source           1..528
                     /organism="Homo sapiens"
                     /db_xref="GDB:5299607"
                     /db_xref="taxon:9606"
                     /clone="A-448E10"
                     /clone_lib="C1978SK"
                     /sex="Female"
                     /cell_type="Fibroblast"
                     /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
                     Caltech Human BAC Library A"

BASE COUNT      118 a 138 c 147 g 125 t
ORIGIN

Query Match      46.5%; Score 67.4; DB 84; Length 528;

/organism="Homo sapiens"
/db_xref="GDB:7147360"
/db_xref="taxon:9606"
/clone="2281N13"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/Note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII"
HindIII
141 a 164 c 166 g 137 t

BASE COUNT      141 a 164 c 166 g 137 t
ORIGIN

Query Match      46.2%; Score 67; DB 84; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGGTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
|||||
DB 34 GTGGATCCCCGGGTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 93
|||||
QY 106 GGGGGCCT 112
|||||
DB 94 GGGGGCCT 100

RESULT 5
A0037147
```

Wed May 31 10:04:55 2000

LOCUS AU037147 141 bp mRNA EST 29-MAR-1999
 DEFINITION AU037147 Dictyostellium discoideum SS (H.Urushihara) Dictyostellium
 discoideum cDNA clone SSB532, mRNA sequence.
 ACCESSION AU037147
 VERSION AU037147.1 GI:3983900
 KEYWORDS EST.
 SOURCE Dictyostellium discoideum.
 ORGANISM Dictyostellium discoideum.
 Eukaryota; Dictyostellida; Dictyostellium.
 REFERENCE 1 (bases 1 to 141)
 AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
 Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takenoto,K., Yasukawa,H.,
 Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
 TITLE The Dictyostellium developmental cDNA project: generation and
 analysis of expressed sequence tags from the first-finger stage of
 development
 JOURNAL DNA Res. 5 (6), 335-340 (1998)
 MEDLINE 99156227
 COMMENT On Jan 19, 1998 this sequence version replaced gi:2153012.
 Contact: Hideo Urushihara
 Institute of Biological Sciences
 University of Tsukuba
 3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
 Email: d402hu@sakura.cc.tsukuba.ac.jp
 PROJECT = 'dictyostellium discoideum cDNA project in Japan'.
 Location/Qualifiers
 1..141
 /organism="Dictyostellium discoideum"
 /strain="Ax4"
 /db_xref="taxon:44689"
 /clone="SSB532"
 /clone_lib="Dictyostellium discoideum SS (H.Urushihara)"
 /dev_stage="slug"
 /dev_stage="slug"
 /dev_stage="slug"

BASE COUNT 26 a 40 c 42 g 28 t 5 others
 ORIGIN

Query Match 46.2%; Score 67; DB 44; Length 141;
 Best Local Similarity 100.0%; Pred. No. 1.4e-12;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGGTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
 |||||||
 Db 69 GTGGATCCCCGGGTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 128

QY 106 GGGGGCC 112
 |||||||
 Db 129 GGGGGCC 135

RESULT 6
 B95279/c 141 bp DNA GSS 26-JUN-1998
 LOCUS CIT-HSP-2172N2.TF CIT-HSP Homo sapiens genomic clone 2172N2,
 DEFINITION genomic survey sequence.
 ACCESSION B95279
 VERSION B95279.1 GI:2977616
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 141)
 AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
 Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
 Simon,M. and Venter,J.C.
 TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
 Building (1998)
 JOURNAL Unpublished (1998)
 COMMENT Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
 Seq primer: M13 Reverse
 Class: BAC ends.
 Location/Qualifiers
 1..144
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="2172N2"
 /clone_lib="CIT-HSP"
 /sex="Male"
 /cell_type="Sperm"
 /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
 HindIII"

BASE COUNT 24 a 41 c 41 g 35 t
 ORIGIN

Query Match 46.2%; Score 67; DB 96; Length 141;
 Best Local Similarity 100.0%; Pred. No. 1.4e-12;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGGTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
 |||||||
 Db 73 GTGGATCCCCGGGTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 14

QY 106 GGGGGCC 112
 |||||||
 Db 13 GGGGGCC 7

RESULT 7
 B80294 144 bp DNA GSS 24-OCT-1998
 LOCUS CIT-HSP-2045D19.TR CIT-HSP Homo sapiens genomic clone 2045D19,
 DEFINITION genomic survey sequence.
 ACCESSION B80294
 VERSION B80294.1 GI:2867317
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 144)
 AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
 Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
 Simon,M. and Venter,J.C.
 TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
 Building (1998)
 JOURNAL Unpublished (1998)
 COMMENT Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
 Seq primer: M13 Reverse
 Class: BAC ends.
 Location/Qualifiers
 1..144
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
 Seq primer: M13-21;
 Class: BAC ends.
 Location/Qualifiers
 1..141
 /organism="Homo sapiens"
 /db_xref="GDB:7103988"
 /db_xref="taxon:9606"
 /clone="2172N2"
 /clone_lib="CIT-HSP"
 /sex="Male"
 /cell_type="Sperm"
 /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
 HindIII"

FEATURES
source

```
/clone="2045D19"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/Note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT      34 a   43 c   41 g   26 t
ORIGIN

Query Match      46.2%; Score 67; DB 84; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
    |||||||
Db 70 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 129
    |||||||
QY 106 GGGGGCC 112
    |||||||
Db 130 GGGGGCC 136

RESULT 8
A0013255/c
LOCUS      A0013255      169 bp      DNA      GSS      06-JUN-1998
DEFINITION CIT-HSP-2299C22.TF CIT-HSP Homo sapiens genomic clone 2299C22,
genomic survey sequence.
ACCESSION  A0013255
VERSION    A0013255.1 GI:3185820
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 169)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
          Building (1998)
JOURNAL   Unpublished (1998)
COMMENT   Other_GSS: CIT-HSP-2299C22.TR
          Contact: Mark Adams
          Department of Eukaryotic Genomics
          The Institute for Genomic Research
          9712 Medical Center Dr., Rockville, MD 20850, USA
          Tel: 301 838 0200
          Fax: 301 838 0208
          Email: mdadams@tigr.org
          Clones are available from Research Genetics (info@resgen.com). BAC
          end search page:
          http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
          Seq primer: M13-21
          Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..169
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="2045D19"
     /clone_lib="CIT-HSP"
     /sex="Male"
     /cell_type="Sperm"
     /note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT      32 a   47 c   48 g   42 t
ORIGIN

Query Match      46.2%; Score 67; DB 96; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.5e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
    |||||||
Db 73 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 132
    |||||||
QY 106 GGGGGCC 112
    |||||||
Db 133 GGGGGCC 139

RESULT 10
T48593
LOCUS      T48593      259 bp      mRNA      EST      02-FEB-1995
DEFINITION phf4_19/1TV Outward Alu-primed hncDNA library Homo sapiens cDNA
clone phf4_19/1TV, mRNA sequence.

Query Match      46.2%; Score 67; DB 96; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.5e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 46 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
    |||||||
Db 101 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 42
    |||||||
QY 106 GGGGGCC 112
    |||||||
Db 41 GGGGGCC 35

RESULT 9
A0041198
LOCUS      A0041198      175 bp      DNA      GSS      14-JUL-1998
DEFINITION CIT-HSP-2335D21.TR CIT-HSP Homo sapiens genomic clone 2335D21,
genomic survey sequence.
ACCESSION  A0041198
VERSION    A0041198.1 GI:3310469
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 175)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
          Building (1998)
JOURNAL   Unpublished (1998)
COMMENT   Contact: Mark Adams
          Department of Eukaryotic Genomics
          The Institute for Genomic Research
          9712 Medical Center Dr., Rockville, MD 20850, USA
          Tel: 301 838 0200
          Fax: 301 838 0208
          Email: mdadams@tigr.org
          Clones are available from Research Genetics (info@resgen.com). BAC
          end search page:
          http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
          Seq primer: M13 Reverse
          Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..175
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="2335D21"
     /clone_lib="CIT-HSP"
     /sex="Male"
     /cell_type="Sperm"
     /note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT      43 a   50 c   47 g   35 t
ORIGIN

Query Match      46.2%; Score 67; DB 96; Length 175;
Best Local Similarity 100.0%; Pred. No. 1.5e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
    |||||||
Db 73 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 132
    |||||||
QY 106 GGGGGCC 112
    |||||||
Db 133 GGGGGCC 139

RESULT 10
T48593
LOCUS      T48593      259 bp      mRNA      EST      02-FEB-1995
DEFINITION phf4_19/1TV Outward Alu-primed hncDNA library Homo sapiens cDNA
clone phf4_19/1TV, mRNA sequence.
```

```

ACCESSION T48593
VERSION T48593.1 GI:642793
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Vinogradova,T.V., Lebedev,Y.B., Kopantzev,E.P., Wagner,L.L., and
JOURNAL Volik,S.V., Ermolaeva,O.D., Lavrentyeva,I., Monastyrskaya,G.S. and
COMMENT Sverdlov,E.D.
OUTWARD Alu-primed hncDNA library
Contact: Sverdlov ED
Structure and Function of Human Genes
Shemyakin Institute of Bioorganic Chemistry
16/10 Mikukho-Makiaya, Moscow, 117871, Russia
Tel: 70953306529
Fax: 70953306538
Email: sverddhumgen.siocb.msk.su.
FEATURES
source
1..259
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="phf14_19/1rv"
/clone_lib="Outward Alu-primed hncDNA library"
/note="Vector: pGEM-3Z; Site_1: EcoRI; Site_2: BamHI; The
library was constructed as described in [Obradovic, D.,
Borodin, A.M., Kopantzev, E.P., Wagner, L.L., Volik, S.V.,
Ermolaeva, O.D., Lebedev, Y.B., Monastyrskaya, G.S.,
Sverdlov, E.D. (1993) Bioorganicheskaya khimiya, 20,
919-930]. This protocol is based on nested primer strategy
using Alu- specific primers (ALN3 and TC-65) that direct
the hncDNA synthesis outward of Alu repeats."
BASE COUNT 62 a 56 c 72 g 59 t
ORIGIN
Query Match 46.2%; Score 67; DB 20; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.6e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
|||||
Db 56 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 115
QY 106 GGGGGCC 122
Db 116 GGGGGCC 122
RESULT 11
AA052885 275 bp mRNA EST 13-SEP-1996
T3551 Bloodstream form of serodeme ILTat1.1 Trypanosoma brucei
brucei cDNA 5', mRNA sequence.
ACCESSION AA052885
VERSION AA052885.1 GI:1543898
KEYWORDS EST.
SOURCE Trypanosoma brucei brucei.
ORGANISM Trypanosoma brucei brucei.
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE 1 (bases 1 to 275)
AUTHORS Osanya,A., Murphy,N.B. and Pelle,R.
TITLE Trypanosoma brucei cDNAs
JOURNAL Unpublished (1996)
COMMENT On Nov 29, 1993 this sequence version replaced gi:637578.
Contact: Osanya A
Unit 3
International Livestock Research Institute
Box 30709, Nairobi, Kenya
Tel: 254 2 630 743
Fax: 254 2 631 499
Email: a.osanya@icgnet.com
Seq primer: M13 primer.
Location/Qualifiers
1..275
/organism="Trypanosoma brucei brucei"
/db_xref="taxon:5702"
/clone_lib="Bloodstream form of serodeme ILTat1.1"
/note="cDNAs were generated from poly (A+) enriched mRNA
prepared from different developmental stages of T.b.brucei
by reverse transcription followed by PCR amplification
using mini-exon and oligo(dT) primers. The cDNA generated
were utilized in RAD5-PCR coupled with differential
hybridisations to identify differentially expressed mRNA
transcripts. The products which showed to be
differentially expressed were cloned pGEM -T vector.
These differentially expressed mRNA transcripts were (are
being) sequenced to generate differentially expressed
sequence tags."
BASE COUNT 61 a 84 c 74 g 56 t
ORIGIN
Query Match 46.2%; Score 67; DB 27; Length 275;
Best Local Similarity 100.0%; Pred. No. 1.6e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 105
|||||
Db 49 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 108
QY 106 GGGGGCC 112
Db 109 GGGGGCC 115
RESULT 12
B74162/c 386 bp DNA GSS 24-OCT-1998
LOCUS CIT-HSP-2028N1.TF CIT-HSP Homo sapiens genomic clone 2028N1,
DEFINITION genomic survey sequence.
ACCESSION B74162
VERSION B74162.1 GI:2769849
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 386)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: CIT-HSP-2028N1.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21
Class: BAC ends.
Location/Qualifiers
1..386
/organism="Homo sapiens"
/db_xref="GDB:7048251"
FEATURES
source
1..386
/organism="Homo sapiens"
/db_xref="GDB:7048251"

```

```

/db_xref="taxon:9606"
/clone="2028N1"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/notes="vector: pBelOBAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT      81 a   98 c   109 g   98 t
ORIGIN

Query Match      46.2%; Score 67; DB 84; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGCTCGACCTCGAGG 105
      |||||||
Db 82 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGCTCGACCTCGAGG 23

Qy 106 GGGGGCC 112
      |||||||
Db 22 GGGGGCC 16

RESULT 13
A0014512/c
LOCUS      A0014512      462 bp      DNA      GSS      06-JUN-1998
DEFINITION CIT-HSP-2300D21.TF CIT-HSP Homo sapiens genomic clone 2300D21,
genomic survey sequence.
ACCESSION  A0014512
VERSION     A0014512.1 GI:3184839
KEYWORDS   GSS.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 462)
AUTHORS     Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE       Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
JOURNAL
COMMENT     Unpublished (1998)
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
            Location/Qualifiers
                1..462
                /organism="Homo sapiens"
                /db_xref="GDB:7154425"
                /db_xref="taxon:9606"
                /clone="2300D21"
                /clone_lib="CIT-HSP"
                /sex="Male"
                /cell_type="Sperm"
                /note="vector: pBelOBAC11; Site_1: HindIII; Site_2:
                HindIII"
BASE COUNT      89 a   127 c   131 g   115 t
ORIGIN

Query Match      46.2%; Score 67; DB 96; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGCTCGACCTCGAGG 105
      |||||||
Db 82 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGCTCGACCTCGAGG 23

Qy 106 GGGGGCC 112
      |||||||
Db 22 GGGGGCC 16

RESULT 15
A0012629
LOCUS      A0012629      547 bp      DNA      GSS      06-JUN-1998
DEFINITION CIT-HSP-2022D24.TF CIT-HSP Homo sapiens genomic clone 2022D24,
genomic survey sequence.
ACCESSION  B65534
VERSION     B65534.1 GI:2639512
KEYWORDS   GSS.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 519)
AUTHORS     Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE       Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
            Unpublished (1997)
            Other_GSSs: CIT-HSP-2022D24.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
            Location/Qualifiers
                1..519
                /organism="Homo sapiens"
                /db_xref="GDB:7045730"
                /db_xref="taxon:9606"
                /clone="2022D24"
                /clone_lib="CIT-HSP"
                /sex="Male"
                /cell_type="Sperm"
                /note="vector: pBelOBAC11; Site_1: HindIII; Site_2:
                HindIII"
BASE COUNT      110 a   137 c   148 g   124 t
ORIGIN

Query Match      46.2%; Score 67; DB 84; Length 519;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGCTCGACCTCGAGG 105
      |||||||
Db 82 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGCTCGACCTCGAGG 23

Qy 106 GGGGGCC 112
      |||||||
Db 22 GGGGGCC 16

RESULT 15
A0012629
LOCUS      A0012629      547 bp      DNA      GSS      06-JUN-1998
DEFINITION CIT-HSP-2022D24.TF CIT-HSP Homo sapiens genomic clone 2022D24,
genomic survey sequence.
ACCESSION  B65534
VERSION     B65534.1 GI:2639512
KEYWORDS   GSS.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 519)
AUTHORS     Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE       Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
            Unpublished (1997)
            Other_GSSs: CIT-HSP-2022D24.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
            Location/Qualifiers
                1..519
                /organism="Homo sapiens"
                /db_xref="GDB:7045730"
                /db_xref="taxon:9606"
                /clone="2022D24"
                /clone_lib="CIT-HSP"
                /sex="Male"
                /cell_type="Sperm"
                /note="vector: pBelOBAC11; Site_1: HindIII; Site_2:
                HindIII"
BASE COUNT      110 a   137 c   148 g   124 t
ORIGIN

```

Wed May 31 10:04:55 2000

DEFINITION CIT-HSP-2299F1.TR CIT-HSP Homo sapiens genomic clone 2299F1,
genomic survey sequence.
ACCESSION AQ012629
VERSION AQ012629.1 GI:3185194
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 547)
REFERENCE Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends

FEATURES
Location/Qualifiers
1..547
/organism="Homo sapiens"
/db_xref="GDB:7154069"
/db_xref="taxon:9606"
/clone="2299F1"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT 154 a 139 c 125 g 129 t
ORIGIN

Query Match 46.2%; Score 67; DB 96; Length 547;
Best Local Similarity 100.0%; Pred. No. 1.8e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 46 GTGGATCCCCGGCTCGAGGATTCGATATCAAGCTTATCGATACCGTCGAGG 105
Db 75 GTGGATCCCCGGCTCGAGGATTCGATATCAAGCTTATCGATACCGTCGAGG 134
Qy 106 GGGGGCC 112
Db 135 GGGGGCC 141

Search completed: May 29, 2000, 21:13:15
Job time: 36514 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:35:29 ; Search time 1214.87 Seconds
(without alignments)
927.949 Million cell updates/sec

Title: us-08-935-377-7
Perfect score: 148
Sequence: 1 GCCCAAAATGAAACATA.....TTGTTTTTGTGGCCCGGCC 148

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, 3808571567 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_bal.*
- 2: gb_ba2.*
- 3: gb_on.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pl1.*
- 8: gb_pl2.*
- 9: gb_pr1.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_sts.*
- 14: gb_sy.*
- 15: gb_un.*
- 16: gb_vi.*
- 17: em_fun.*
- 18: em_hum1.*
- 19: em_hum2.*
- 20: em_in.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_sy.*
- 29: em_un.*
- 30: em_vl.*
- 31: gb_htg1.*
- 32: gb_htg2.*
- 33: gb_in1.*
- 34: gb_in2.*
- 35: gb_in3.*
- 36: em_bal.*
- 37: em_ba2.*
- 38: em_hum3.*
- 39: em_hum4.*
- 40: gb_pr4.*
- 41: gb_htg3.*
- 42: gb_htg4.*
- 43: gb_htg5.*

Result No.	Score	Query Match %	Length	DB	ID	Description
c 1	59.8	47.2	259	5	A44281	A44281 Sequence 9
c 2	69.4	46.9	2958	14	ARBLKSM	X52326 pBluescript
c 3	69.4	46.9	2958	14	ARBLKSP	X52331 pBluescript
c 4	69.4	46.9	2961	14	ARBL2KSM	X52339 pBluescript
c 5	69.4	46.9	2961	14	ARBL2KSP	X52337 pBluescript
c 6	69.4	46.9	2964	14	SYNBLKSMV	L08784 Bluescribe
c 7	69.4	46.9	2964	14	SYNBLKSPV	L08785 Bluescribe
c 8	69.4	46.9	3306	14	SYNBPEN66	D85525 cloning vec
c 9	69.4	46.9	4144	14	XXU35131	U35131 plasmid pBS
c 10	69.4	46.9	4267	14	PRS304	U35136 Yeast integ
c 11	69.4	46.9	4289	14	XXU35136	U35136 plasmid pBS
c 12	69.4	46.9	4373	14	PRS306	U35138 Yeast integ
c 13	69.4	46.9	4443	14	PRS303	U35135 Yeast integ
c 14	69.4	46.9	4670	14	ASAJ5326	AJ005326 pGAI(++)
c 15	69.4	46.9	4670	14	ASAJ5329	AJ005329 pGAI(-)
c 16	69.4	46.9	4707	14	XXU02374	U02374 Cloning vec
c 17	69.4	46.9	4768	14	XXU25061	U25061 Cloning vec
c 18	69.4	46.9	4783	14	PRS314	U03440 Yeast centr
c 19	69.4	46.9	4887	14	PRS316	U03442 Yeast centr
c 20	69.4	46.9	4950	14	XXU25060	U25060 Cloning vec
c 21	69.4	46.9	4967	14	PRS313	U03439 Yeast centr
c 22	69.4	46.9	5144	14	CVU23751	U23751 Cloning vec
c 23	69.4	46.9	5187	14	U34887	U34887 Yeast integ
c 24	69.4	46.9	5228	14	XXU25059	U25059 Cloning vec
c 25	69.4	46.9	5504	14	PRS305	U03437 Yeast integ
c 26	69.4	46.9	5634	14	CVU14125	U14125 Cloning vec
c 27	69.4	46.9	6018	14	PRS315	U03441 Yeast centr
c 28	69.4	46.9	6340	14	ASAJ5323	AD005323 pCP1(-) K
c 29	69.4	46.9	9655	14	SYNPR8V	AD001531 Cloning v
c 30	68.8	46.5	2961	14	CVU46017	U46017 Cloning vec
c 31	68.8	46.5	3240	8	AF015771	AF015771 Magnaport
c 32	68.8	46.5	3399	14	CVU46018	U46018 Cloning vec
c 33	68.8	46.5	4278	14	AF028239	AF028239 Mammalian
c 34	68.8	46.5	4307	14	AF067646	AF067646 Cloning v
c 35	68.6	46.4	4328	14	AF072999	AF072999 Cloning v
c 36	68.4	46.2	7933	5	AR060142	AR060142 Sequence
c 37	68.4	46.2	4323	14	AF072538	AF072538 Cloning v
c 38	68.4	46.2	4324	14	AF072539	AF072539 Cloning v
c 39	68.4	46.2	4325	14	AF072540	AF072540 Cloning v
c 40	68.4	46.2	4326	14	AF072997	AF072997 Cloning v
c 41	68.4	46.2	4327	14	AF072998	AF072998 Cloning v
c 42	68.2	46.1	3727	14	SYNPROCL	M32616 Synthetic c
c 43	68	45.5	338	1	BSPX91477	X91477 Bacterial s
c 44	68	45.5	6037	14	YSCTRAM2	M74016 Cloning vec
c 45	68	45.5	6445	14	SYNECOYST	L11060 Cloning vec

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS


```

#parent
  VecBase(3):BlueM13m
#parent
  VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
  VecSource(3):bGalks, GenBank(50):PFI
#brother
  VecBase(3):BlueKsp, VecBase(3):BlueKsm
PARENT
  Features of BlueKsm (2964 bp)
    residue source
    3- 458 5488-5943 phage f1
    460- 624 236- 400 pUC19c
    626- 645 1- 20 T7 promoter
    653- 760 108- 1 (c) BlueKs-polylinker
    772- 791 20- 1 (c) T3 promoter
    795-2964 448-2617 pUC19c
    Conflict (cfl) and Mutations (mut): none
PARENT
  Features of BlueKsm (2964 bp)
    residue source
    3- 458 5488-5943 phage f1
    460- 624 449- 285 (c) pUC19
    626- 645 1- 20 T7 promoter
    653- 760 108- 1 (c) BlueKs-polylinker
    772- 791 20- 1 (c) T3 promoter
    795-1031 237- 1 (c) pUC19
    1032-2964 2686- 754 (c) pUC19
    Conflict (cfl) and Mutations (mut): none
FEATURE
  643 start of T7-RNA synthesis
  774 (c) start of T3-RNA synthesis
  1976-2764 789-1 (c) Ap-R; b-lactamase
POLYLINKER
  KpnI-DraII-ApaI-XhoI-SalI-ClaiI-HindIII-EcoRV-EcoRI-PstI-
  SmaI-BamHI-SpeI-XbaI-NotI-XmaII-BstXI-SacII-SELECTION
#resistance Ap
#indicator beta-galactosidase
#length 2964 #checksum 1589.
SUMMARY BlueKsm
  Location/Qualifiers
  1..2964
  /organism="synthetic construct"
  /db_xref="taxon:32630" 765 t
BASE COUNT 708 a 756 c 735 g 765 t
ORIGIN

Query Match 46.9%; Score 69.4; DB 14; Length 2964;
Best Local Similarity 87.4%; Pred. NO. 1.6e-11;
Matches 76; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 47 TGGTGGATCCCGGGCTGCAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
Db 685 TAGTGGATCCCGGGCTGCAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 744
QY 107 GGGGGGGCCCTAACTAACTTTTGT 133
Db 745 GGGGGGGCCGTTACCCAGCTTTTGT 771

RESULT 7
SYNBLKSPV SYNBLKSPV 2964 bp DNA circular SYN 26-JUL-1993
LOCUS Bluescribe KS Plus cloning vector.
DEFINITION L08785
ACCESSION L08785
VERSION L08785.1 GI:310729
KEYWORDS Synthetic construct DNA.
ORGANISM synthetic construct
          artificial sequence.
REFERENCE 1 (bases 1 to 2964)
AUTHORS Gilbert W.
TITLE Obtained from Vecbase 3.0
JOURNAL Unpublished (1991)

```

COMMENT

```

These data and their annotation were supplied to GenBank by Will
Gilbert under the auspices of the GenBank Curator Program.
Bluescribe KS Plus - Cloning vector
ENTRY BLUEKSP #TYPE DNA CIRCULAR
TITLE Bluescribe KS Plus - Cloning vector
DATE 28-JAN-1987
#sequence 02-FEB-1987
#sequence 04-MAR-1987
#sequence 03-APR-1987
ACCESSION VB0078
SOURCE artificial
REFERENCE
#number 1
#authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
J.
#journal Gene (1987) in press
#citation Sequence data from Stratagene
#comment Sequence correction according to Stratagene COMMENT
Obtained from Stratagene on floppy disc.
Revised 02-FEB-1987 by F. Pfeiffer:
1409/10 'AT' to 'TA' to match revised sequence of PBR322
Revised 4-MAR-1987 to match sequence of pUC19 on request
of Stratagene
Polylinker region revised 03-APR-1987 according to Stratagene
COMMENT
The stand shown corresponds to pUC19c.
As in the published sequence of pUC19c, The M13mp19 lacZ region
is on the complementary strand.
COMMENT
This vector contains the f1 origin so that the plus strand
can be obtained upon f1 superinfection.
KEYWORDS
CROSSREFERENCE
#parent
  VecBase(3):BlueM13p
#parent
  VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
  VecSource(3):bGalks, GenBank(50):PFI
#brother
  VecBase(3):BlueKsm, VecBase(3):BlueKsp
PARENT
  Features of BlueKsp (2964 bp)
    residue source
    3- 458 5943-5488 (c) phage f1
    460- 624 236- 400 pUC19c
    626- 645 1- 20 T7 promoter
    653- 760 108- 1 (c) BlueKs-polylinker
    772- 791 20- 1 (c) T3 promoter
    795-2964 448-2617 pUC19c
    Conflict (cfl) and Mutations (mut): none
PARENT
  Features of BlueKsp (2964 bp)
    residue source
    3- 458 5943-5488 (c) phage f1
    460- 624 449- 285 (c) pUC19
    626- 645 1- 20 T7 promoter
    653- 760 108- 1 (c) BlueKs-polylinker
    772- 791 20- 1 (c) T3 promoter
    795-1031 237- 1 (c) pUC19
    1032-2964 2686- 754 (c) pUC19
    Conflict (cfl) and Mutations (mut): none
FEATURE
  643 start of T7-RNA synthesis
  774 (c) start of T3-RNA synthesis
  1976-2764 789-1 (c) Ap-R; b-lactamase
POLYLINKER
  KpnI-DraII-ApaI-XhoI-SalI-ClaiI-HindIII-EcoRV-EcoRI-PstI-
  SmaI-BamHI-SpeI-XbaI-NotI-XmaII-BstXI-SacII-SELECTION
#resistance Ap
#indicator beta-galactosidase
#length 2964 #checksum 690.
SUMMARY BlueKsp
  Location/Qualifiers
  1..2964
  /organism="synthetic construct"
  /db_xref="taxon:32630" 765 t
BASE COUNT 708 a 756 c 735 g 765 t
ORIGIN

```



```
REFERENCE 1 (bases 1 to 4373)
AUTHORS Sikorski,R.S. and Hietter,P.
TITLE A system of shuttle vectors and yeast host strains designed for
efficient manipulation of DNA in Saccharomyces cerevisiae
JOURNAL Genetics 122 (1), 19-27 (1989)
MEDLINE 89276910
REFERENCE 2 (bases 1 to 4373)
AUTHORS Stillman,D.J.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
and Molecular Biology, University of Utah Medical Center, Salt Lake
City, UT 84132, USA
FEATURES
source
1. .4373
/organism="Cloning vector PRS306"
/db_xref="taxon:31829"
BASE COUNT 1181 a 987 c 1106 g 1099 t
ORIGIN
Query Match 46.9%; Score 69.4; DB 14; Length 4373;
Best Local Similarity 87.4%; Pred. No. 1.7e-11;
Matches 76; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 47 TGTGTGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
Db 2025 TAGTGTGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 2084
QY 107 GGGGGGGCCTAACTAACTAATTTTGT 133
Db 2085 GGGGGGGCCGGTACCCAGCTTTTGT 2111
RESULT 13
PRS303 4443 bp DNA circular SYN 14-SEP-1995
LOCUS Yeast integrative vector PRS303 with HIS3 marker, complete
DEFINITION sequence.
ACCESSION U03435
VERSION U03435.1 GI:416304
KEYWORDS Cloning vector PRS303.
SOURCE Cloning vector PRS303.
ORGANISM artificial sequence; vectors.
REFERENCE 1 (bases 1 to 4443)
AUTHORS Sikorski,R.S. and Hietter,P.
TITLE A system of shuttle vectors and yeast host strains designed for
efficient manipulation of DNA in Saccharomyces cerevisiae
JOURNAL Genetics 122 (1), 19-27 (1989)
MEDLINE 89276910
REFERENCE 2 (bases 1 to 4443)
AUTHORS Stillman,D.J.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
and Molecular Biology, University of Utah Medical Center, Salt Lake
City, UT 84132, USA
FEATURES
source
1. .4443
/organism="Cloning vector PRS303"
/db_xref="taxon:31826"
BASE COUNT 1149 a 1048 c 1109 g 1137 t
ORIGIN
Query Match 46.9%; Score 69.4; DB 14; Length 4443;
Best Local Similarity 87.4%; Pred. No. 1.7e-11;
Matches 76; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 47 TGTGTGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
Db 2105 TAGTGTGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 2164
QY 107 GGGGGGGCCTAACTAACTAATTTTGT 133
```

```
Db 2165 GGGGGGGCCGGTACCCAGCTTTTGT 2191
RESULT 14
ASAJ5326/c 4670 bp DNA circular SYN 08-FEB-1999
LOCUS pGATII(+) KS positive selection cloning vector glts gene.
DEFINITION AJ005326
ACCESSION AJ005326
VERSION AJ005326.1 GI:4028984
KEYWORDS glts gene; glutamate permease.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 4670)
AUTHORS Gal,J.
TITLE Direct Submission
JOURNAL Submitted (03-APR-1998) Gal J., Institute for Biotechnology, Bay
Zoltan Foundation for Applied Research, Szeged, Derkovits fasor 2.,
6726, HUNGARY
REFERENCE 2 (bases 1 to 4670)
AUTHORS Gal,J., Szekeres,S., Schnell,R., Pongor,S., Simoncsits,A. and
Kalman,M.
TITLE A positive selection cloning system based on the glts gene of
Escherichia coli
JOURNAL Anal. Biochem. 266 (2), 235-238 (1999)
MEDLINE 99107575
FEATURES
source
1. 4670
/organism="synthetic construct"
/db_xref="taxon:32630"
852. .2231
/gene="glts"
852. .2231
/gene="glts"
/codon_start=1
/transl_table=11
/product="glutamate permease"
/protein_id="CAA06473.1"
/db_xref="GI:4028985"
/translation="MFHLDTLVAATLLLRKLVHSYFLKKYTPPEPVAGLL
VALALLVKKSMGEVNEFDMSLRDLPLMLAFATIGLNANIASLRAGRVGFIILIV
GILVQNAIGIGMASLLGLDPLMLAGLISILSGHGCGAASKLFIERYGFTNATIEV
AMACATFLGVLGGIGPVARQLTKGNKSWPPSPSTVSISLNSCSPGDLV
LERPPRWSNSPYSESYARLYVKHSTPNIPDDQEVPTAFERPDYGRMTISLVI
ETIALIAICLVAGKIVAOAGLAFELPFCVLFVGVILSGLSIMGFYRFRERAVS
VLGNVLSLFLAMALMGILKELASLALPLMLAILVQVIFMALYAFVYWRMMGKNYD
RAVLAAGHCGFLGATPTAIANMQAITERFGPSHMAFLVPMVGAFDIDVNALVIKL
YLMLPFAG"
BASE COUNT 1040 a 1165 c 1231 g 1234 t
ORIGIN
Query Match 46.9%; Score 69.4; DB 14; Length 4670;
Best Local Similarity 87.4%; Pred. No. 1.7e-11;
Matches 76; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 47 TGTGTGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
Db 1502 TAGTGTGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 1443
QY 107 GGGGGGGCCTAACTAACTAATTTTGT 133
Db 1442 GGGGGGGCCGGTACCCAGCTTTTGT 1416
RESULT 15
ASAJ5329
LOCUS pGATII(-) KS positive selection cloning vector glts gene.
DEFINITION AJ005329
ACCESSION AJ005329
VERSION AJ005329.1 GI:4028990
KEYWORDS glts gene; glutamate permease.
```

SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 4670)
AUTHORS Gal, J.
TITLE Direct Submission
JOURNAL Submitted (03-APR-1998) Gal J., Institute for Biotechnology, Bay
Zoltan Foundation for Applied Research, Szeged, Derkovits faszor 2.,
6726, HUNGARY
REFERENCE 2 (bases 1 to 4670)
AUTHORS Gal, J., Szekeres, S., Schnell, R., Pongor, S., Simoncsits, A. and
Kalman, M.
TITLE A positive selection cloning system based on the glts gene of
JOURNAL Escherichia coli
MEDLINE Anal. Biochem. 266 (2), 235-238 (1999)
FEATURES 99107575 Location/Qualifiers
 1..4670
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 complement(1150..2529)
 /gene="glts"
 complement(1150..2529)
 /gene="glts"
 /codon_start=1
 /transl_table=11
 /product="glutamate permease"
 /protein_id="CAA06476.1"
 /db_xref="GI:4028991"
 /translation="MFHLDTLATLVAAATLLGLGRKLVHSVFLKKYTIPEPVAGLL
 VALALLVKKSMGWEVDFMSLRDPLMLAFATIGLNANIASLRAGRVVGIPLIVV
 GLLVQNAIGIGMALGLDPLMGLLAGSITLGGHGTGAWSKLFIERYGFTNATEV
 AMACATGLVIGGIIGPVARQLTLTKGNKSWVPSPSTVSIISNSCPDPLV
 LERPPRWSNSPISEYARYLVKHSTPNGIPDDQEVPTAFKPDVGRMITSVLVLI
 ETIALAICLVTVGRIVAQLLAGTAFELPTFVCLFVGLISNGLSIMGFYRPERAVS
 VLGNVSLFLFAMALMGLKWLWELASLAPLAILLVQTIIFMALYAFVTVWRMGKNYD
 AAVLAAGHCSCGFLGATPTATANMQAITERFGPSHMAFLVPMVGAFFIDIVNALVINK
 YLMPIFAG"
BASE COUNT 1198 a 1251 c 1145 g 1076 t
ORIGIN

Query Match 46.9%; Score 69.4; DB 14; Length 4670;
Best Local Similarity 87.4%; Pred. No. 1.7e-11;
Matches 76; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 47 TGGTGGATCCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
 |||||
Db 1879 TAGTGGATCCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 1938
 |||||
QY 107 GGGGGGGCCCTAACTAACTAATTTGTT 133
 |||||
Db 1939 GGGGGGGCCCGGTACCCAGCCTTTGTT 1965
 |||||

Search completed: May 29, 2000, 21:35:34
Job time: 36866 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:58:21 ; Search time 1446.77 Seconds
(without alignments)
25.594 Million cell updates/sec

Title: US-08-935-377-7
Perfect score: 148
Sequence: 1 GGCCAAATGAAACTA.....TTGTTTTGTGGCGCCGCC 148

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	71.8	48.5	10529	1 V09028	Maize fluory2 gene
C 2	69.8	47.2	259	1 Q87664	Mouse azoospermia
C 3	69.8	47.2	501	1 Q87666	Nucleotide analogu
C 4	68.6	46.4	1091	1 T75006	Human endogenous r
C 5	68.6	45.9	5356	1 T43794	Plasmid pRIPRAT (r
C 6	67.6	45.7	1122	1 T75005	Human endogenous r
C 7	67.4	45.5	651	1 X20513	Polynucleotide seq
C 8	67.4	45.5	752	1 V31294	E. coli J96 pathog
C 9	67.4	45.5	3792	1 Q48463	Plasmid pg+host4 c
C 10	67.4	45.5	4226	1 T39485	Human steroidogene
C 11	67.4	45.5	5234	1 Q48464	Plasmid pg+host5 c
C 12	67.4	45.5	6722	1 Q48465	Plasmid pg+host6 c
C 13	67.4	45.5	12814	1 X24730	Swedish-FAD APP ta
C 14	67.4	45.5	15892	1 X24731	London-FAD APP tar
C 15	67.4	45.5	15892	1 X24732	Swedish/London-FAD
C 16	67.4	45.5	15701	1 X24733	Swedish-FAD APP713
C 17	67.2	45.4	545	1 V68808	Human endogenous r
C 18	65.4	44.2	2640	1 Q26664	bDAT, cDNA encodin
C 19	65	43.9	84	1 X02800	E. coli biotin DNA
C 20	65	43.9	3465	1 X02814	DE19731274 Seq ID
C 21	65	43.9	3481	1 X02815	DE19731274 Seq ID
C 22	65	43.9	3794	1 X02812	DE19731274 Seq ID
C 23	65	43.9	3810	1 X02813	DE19731274 Seq ID
C 24	65	43.9	6596	1 V57377	Maize female-prefe
C 25	64	43.2	545	1 T75010	Human endogenous r
C 26	62.8	42.4	685	1 V10190	Stealth virus nucl
C 27	62.8	42.4	685	1 V12003	Stealth virus plas
C 28	59.6	40.3	4164	1 T04575	Plasmid pAT-1 sequ
C 29	59.6	40.3	4164	1 V22271	pAT-1 (pSD544). DN
C 30	59.6	40.3	4164	1 V69740	Nucleotide sequenc
C 31	58.4	39.5	5178	1 T43876	pret-Splice. Nucle
C 32	58.2	39.3	3198	1 T92702	Candida Carhol gen
C 33	58.2	39.3	3198	1 T92869	Candida Carhol gen

RESULT 1

V09028/c

ID V09028 standard; DNA; 10529 BP.

AC V09028;

DT 20-JUL-1998 (first entry)

DE Maize fluory2 gene (fl2).

KW Flury2; fl2 gene; alpha-zein; signal peptide; transgenic plant;

KW seed; essential amino acid; animal feed; maize; rice; wheat;

KW barley; millet; sorghum; ds.

OS zea mays.

FH Key Location/Qualifiers

FT misc_feature 1..760

FT /tag= a

FT /note= "vector-derived sequence"

FT promoter 761..3824

FT /tag= b

FT /note= "Claim 14"

FT CDS 3825..4613

FT /tag= c

FT sig_peptide 3825..3887

FT /tag= d

FT mat_peptide 3888..4610

FT /tag= e

FT W09802563-A1.

PN 22-JAN-1998.

PD 11-JUL-1997; U11723.

PF 17-JUL-1996; US-021833.

PR (PION-) PIONEER HI-BRED INT INC.

PA (UYAR-) UNIV ARIZONA STATE.

PI Beach L, Coleman CE, Larkins BA;

PL WPI; 98-110609/10.

DR P-PSDB; W23977.

CC Cereal plants containing trans-gene expressing fusion that includes

CC signal peptide of the fl2 maize gene - and protein having high

CC content of essential amino acids, producing feeds of improved

CC nutritional value

PS Claim 14; Fig 1A-H; 37pp; English.

CC This is the nucleotide sequence of a clone of the fluory2 (fl2)

CC gene of maize. It codes for a 24-kDa alpha-zein protein (see

CC W23977). that includes a 21-amino acid signal peptide (see W23976)

CC which targets the alpha-zein to the lumen of the rough endoplasmic

CC reticulum. A claimed cereal plant contains a transgene comprising

CC a first polynucleotide that encodes the fl2 signal peptide and a

CC second polynucleotide that encodes an agronomically high-value

CC protein. Also new are seeds produced by the plants, the transgene

CC itself, and a transgene that also includes the fl2 promoter. The

CC second polynucleotide preferably encodes a protein that has a high

CC content of Met, Lys, Trp and/or Thr so that feeds from transformed

CC maize, wheat, rice, barley, millet or sorghum will have increased

CC contents of these essential amino acids in their seeds (all

CC claimed).

SQ Sequence 10529 BP; 2845 A; 2484 C; 2262 G; 2938 T;

Query Match 48.5%; Score 71.8; DB 1; Length 10529;
Best Local Similarity 76.5%; Pred. NO. 9.8e-16;

Clone #4 from muta
Nucleotide sequenc
E. coli biotin DNA
Yeast checkpoint c
Yeast RAD17 coding
Plasmid pK5/14 D
Plasmid pK5/6 DNA
Plasmid pK9/10 DN
Plasmid pK7/8 DNA
DNA sequence of ex
S. aureus Murf ORF

Matches	88;	Conservative	0;	Mismatches	27;	Indels	0;	Gaps	0;
QY	1	GGCAAAAATTTGA	AAACTAGATCTATTTATTGCACGGGGCGCCATGGTGGATCCCGG	60					
DB	774	GGCTAAATTTGCAT	GAAGCTCCACGGGGTGGCGGCTCTAGAACTASTAGATCCCGG	715					
QY	61	GGCTGCAGGAATTC	GCATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC	115					
DB	714	GGCTGCAGGAATTC	GCATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCC	660					
RESULT	2								
Q87664/c									
ID	Q87664	standard; DNA; 259 BP.							
AC	Q87664;								
DT	06-MAR-1996	(first entry)							
DE	Mouse azoospermia factor (AZF)	gene partial clone M3.2.							
KW	Azoospermia factor; AZF; male infertility; YRRM gene;								
OS	Y-chromosome; RNA recognition motif; ds.								
Key	Mus sp.								
FT	Key	Location/Qualifiers							
FT	misc_difference 174								
FT	/*tag= a								
FT	/note= "base n at position 174 is not identified								
FT	in the specification"								
FT	misc_difference 197								
FT	/*tag= b								
FT	/note= "base n at position 197 is not identified								
FT	in the specification"								
PN	W09511300-A2.								
PD	27-APR-1995.								
PF	24-OCT-1994; G02344.								
PR	22-OCT-1993; GB-021857.								
PR	07-JUL-1994; GB-013760.								
PA	(MEDI-) MEDICAL RES COUNCIL.								
PI	Chandley AC, Cooke HJ, Hargreave TB, Kun M, Sharkey AM;								
PT	WPI: 95-170221/22.								
PT	Nucleic acid encoding the human azoospermia factor, and probes and								
PT	antibodies specific for the sequence and encoded polypeptide - may								
PT	be used in the clinical diagnosis of male infertility								
PS	Disclosure: Fig 5; 40pp; English.								
CC	A mouse gene partial sequence (Q87664) showing homology to the human								
CC	azoospermia factor (AZF) YRRM gene was obtd. from mouse genomic phage								
CC	clone M3.2 isolated from a genomic library screened with MYK (YRRM1)								
CC	CDNA (Q87655). M3.2, cloned in Lambda 2001, was mapped to the short arm								
CC	of the mouse Y-chromosome. A full cDNA sequence (Q87665) was also								
CC	obtd.								
SC	Sequence 259 BP; 67 A; 58 C; 63 G; 69 T;								
Query Match	47.2%;	Score 69.8; DB 1; Length 259;							
Best Local Similarity	79.0%;	Pred. No. 1.4e-15;							
Matches	83; Conservative	0; Mismatches 22; Indels 0; Gaps 0;							
QY	29	TATTGACGGGCGCCATCGTGGATCCCGGGCTCGAGGAATTCGATATCAAGCTTAT	88						
DB	132	TAAAGAAAGCAAGCCCTCTAGTGGATCCCGGGCTCGAGGAATTCGATATCAAGCTTAT	73						
QY	89	CGATACGCTGCACCTCGAGGGGGCGCTAACTAATTTTGT	133						
DB	72	CGATACGCTGCACCTCGAGGGGGCGCTAATTTTGT	28						
RESULT	3								
T04866									
ID	T04866	standard; DNA; 501 BP.							
AC	T04866;								
DT	28-JAN-1996	(first entry)							
DE	Nucleotide analogue treated with calf intestinal alkaline phosphatase.								
KW	Nucleotide analogue; alkaline phosphatase; ss.								
OS	Synthetic.								
Key	Key	Location/Qualifiers							
FT	misc feature 1. .501								

	/tag=	a	
	/note=-	"std. IUPAC codes used"	
WT	W09524505-A.		
PV	14-SEP-1995.		
PD	07-MAR-1995:	G00490.	
PF	08-MAR-1994:	EP-301636.	
PR	(AMSH) AMERSHAM INT PLC.		
PA	Reeve MA,	Robinson PS;	
PI	WPJ; 95-328290/42.		
DR	Modification of residual fluorescence labelled nucleotide analogues - to prevent migration in electrophoretic sequencing gel and interference with base calling of DNA chains		
PT	Example: Fig 1; lbp; English.		
CC	The invention concerns the modification of redial fluorescently labelled nt analogues to prevent migration in electrophoretic sequencing gel and interference with base calling of DNA chains. The modification involves the use of a phosphatase enzyme to remove at least one 5'-phosphate gp. The nts were prepared using ABI Amplitaq dye-terminating kt. The template used was 1 microg. Of M13 mp8. The primer was M13 Universal primer. Samples were recovered. Some were subjected to calf intestinal alkaline phosphatase digestion. This includes R04866. T04867 was a control nt which was not treated with alkaline phosphatase. The example shows that alkaline phosphatase treatment causes removal of the dye-terminator artefacts and allows for accurate base calling with the ABI analysis software.		
CC	Sequence	501 BP;	98 A; 136 C; 140 G; 121 T;
SQ			
	Query Match	47.2%; Score 69.8; DB 1; Length 501;	
	Best Local Similarity	81.6%; Pred. No. 1.7e-15;	
	Matches	80; Conservative	0; Mismatches 18; Indels 0; Gaps 0;
QY	36 CGGGGCCCGCATGTGGATCCCCGCGCTGCAGGAATTTCGATATCAAGTTTTCGATAACC	95 	
Dd	44 GCNC TAGA ACTAGTGGA TCCTCCGCGGCTGCAGGAATTTCGATATCAAGTTTTCGATAACC	103 	
QY	96 GTCCACTC TGAGGGGGGCGCTTA CTAACTAA TTGTT	133 	
Dd	104 GTCCACTC TGAGGGGGGCGCGGT ACCAGCTTTTGT	141 	
RESULT	4		
T75006/c	ID T75006 standard; DNP; 1091 BP.		
AD	I75006;		
DC	06-OCT-1997 (first entry)		
DT	Human endogenous retroviral sequence 6.		
DE	Breast cancer; tumor; B18Ag1; prognosis; diagnosis; vaccine; ps.		
KW	Human retrovirus.		
OS	WO9725431-Al.		
PN	17-JUL-1997.		
PD	10-JAN-1997; U00398.		
PR	10-JAN-1996; US-587329.		
PA	(CORI-) CORIXA CORP.		
FI	Frudakis TN, Smith JM;		
DR	WPI; 97-384982/35.		
PT	Endogenous human tumour-associated retroviral element, B18Ag1 - used for the prognosis, diagnosis and monitoring of human cancers,		
PT	especially breast cancer		
PS	Claim 10; Page 31-32; 74pp; English.		
CC	Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14		
CC	and 11-22 (T75003-10) were obtd by screening human genomic libraries using human breast tumour-associated retroviral element B18Ag1 (see also T75002) as probe. These non-contiguous sequences lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the CC retrovirus genome (see also T75001). B18Ag1 and the other CC retroviral sequences can be used in genetic vaccines and for the CC prognosis, diagnosis and monitoring of human breast cancer.		
SQ	Sequence	1091 BP;	79 A; 350 C; 97 G; 248 T;
	Query Match	46.4%; Score 68.6; DB 1; Length 1091;	
	Best Local Similarity	80.8%; Pred. NO. 6e-15;	

Db	743	TCACCTTGGATGGTTGGGGATCCCGGGGCTGCAGGAATTCGATATCAAGCTTATCGA	660
QY	92	TACCCTCGACCTCGAGGGGGGCC	115
Db	683	TACCCTCGACCTCGAGGGGGGCC	660
RESULT	6		
ID	T75005/c		
IT	T75005 standard; DNA; 1122 BP.		
AC	T75005:		
DT	06-OCT-1997 (first entry)		
DE	Human endogenous retroviral sequence 3.		
KS	Breast cancer; tumour; B18Ag1; prognosis; diagnosis; vaccine; ss.		
OW	Human retrovirus.		
PN	WO9725431-A1.		
PD	17-JUL-1997.		
PF	10-JAN-1997; U00398		
PR	10-JAN-1996; US-587329.		
PA	(CORI-) CORIXA CORP.		
PI	Frudakis TN, Smith JM;		
WPI	97-384982/35.		
DR	Endogenous human tumour-associated retroviral element, B18Ag1 - used		
PT	for the prognosis, diagnosis and monitoring of human cancers,		
PT	especially breast cancer		
PS	Claim 10; Page 30-31; 74pp: English.		
CC	Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14		
CC	and 11-22 (T75003-10) were obtd. by screening human genomic		
CC	libraries using human breast tumour-associated retroviral element		
CC	B18Ag1 (see also T75002) as probe. These non-contiguous sequences		
CC	lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the		
CC	retrovirus genome (see also T75001). B18Ag1 and the other		
CC	retroviral sequences can be used in genetic vaccines and for the		
CC	prognosis, diagnosis and monitoring of human breast cancer.		
CC	Sequence 1122 BP; 260 A; 316 C; 185 G; 279 T;		
SEQ			
Query Match	45.7%; Score 67.6; DB 1; Length 1122;		
Best Local Similarity	80.6%; Pred. No. 1.4e-14;		
Matches 79; Conservative	0; Mismatches 19; Indels 0; Gaps		
QY	18 CTAGATCATTTATTCGACGGCGCCAGTCGTTGGATCCCGGGCTGCAGGAATTCGAT	77	
Db	139 CTAAATCTCTCATTTCTTGTCTAGAACATAGTGGATCCCGGGCTGCAGGAATTCGAT	80	
QY	78 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCC	115	
Db	79 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCC	42	
RESULT	7		
ID	X20513		
AC	X20513 standard; DNA; 651 BP.		
DT	05-MAY-1999 (first entry)		
DE	Polynucleotide sequence from the genome of Treponema pallidum.		
KS	Treponema pallidum infection; syphilis; Borrelia infection; animal;		
KW	enzyme production; ds.		
PN	Treponema pallidum.		
OW	WO9859034-A2.		
PD	30-DEC-1998.		
PF	23-JUN-1998; U13041.		
PR	24-JUN-1997; US-050667.		
PA	(HUNA-) HUMAN GENOME SCI INC.		
PI	Fraser CM;		
WPI	99-081273/07.		
DR	New isolated Treponema pallidum nucleic acids - used to develop		
PT	products for the detection, diagnosis, characterisation, prevention		
PT	and therapy of T. pallidum infections, particularly syphilis		
PS	Claim 1; Page 257; 1150pp: English.		
CC	X20500-21243 represent polynucleotide sequences from the genome of		
CC	Treponema pallidum. The sequences can be used for detection,		
CC	diagnosis, characterisation, prevention and therapy for. pallidum		

RESULT	9
Q48463/C	
ID	Q48463 standard; DNA; 3792 BP.
AC	Q48463;

PN WO6293338-AL.
PD 26-SEP-1996.
PE 22-MAR-1996; U03896.
PF 23-MAR-1995; US-410540.
PG PR (REG) UNIV CALIFORNIA.
PH PA (UYPE-) UNIV PENNSYLVANIA.
PI Lin D, Miller WL, Strauss JF;
PJ WPI: 96-443130/44.
PK Isolated human steroidogenesis acute regulatory protein gene - used
PL for detection of mutations(s) of this gene that cause congenital
PM lipid adrenal hyperplasia
PN Claim 1; Pages 23-25; 89pp; English.
PO The present sequence encodes the human steroidogenesis acute
PP regulatory protein (hSTAR). The hSTAR gene can be analysed for
PQ mutations to detect (e.g. prenatally) genetic defects associated
PR with congenital lipid adrenal hyperplasia (CAH), or its
PS transmutation to children. CAH can be treated by protein or gene
PT replacement therapy, which can also be used to prevent or treat
PU hypercholesterolaemia.
PV
PW
PX
PY
PZ

2

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 22:09:00 ; Search time 621.83 Seconds
(without alignments)
30.937 Million cell updates/sec

Title: us-08-935-377-7
Perfect score: 148
Sequence: 1 GGCAAAATGAAACATA.....TTCTTTTGTGGCCGGCC 148

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgnl_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgnl_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgnl_6/ptodata/1/ina/5C_COMB.seq.*
4: /cgnl_6/ptodata/1/ina/5D_COMB.seq.*
5: /cgnl_6/ptodata/1/ina/6_COMB.seq.*
6: /cgnl_6/ptodata/1/ina/PCTUS_COMB.seq.*
7: /cgnl_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	68.4	46.2	793	3	US-08-967-101-121
C 2	68.4	46.2	793	4	US-08-592-541-116
C 3	67.4	45.5	3792	4	US-08-992-334-1
C 4	67.4	45.5	3792	5	US-08-302-752-1
C 5	67.4	45.5	4016	2	US-08-410-540-3
C 6	67.4	45.5	5234	4	US-08-992-334-2
C 7	67.4	45.5	5234	5	US-08-302-752-2
C 8	67.4	45.5	6722	4	US-08-992-334-3
C 9	67.4	45.5	6722	5	US-08-302-752-3
C 10	65.4	44.2	9318	5	US-08-793-610-6
C 11	62.8	42.4	685	2	US-08-463-115-56
C 12	62.8	42.4	685	2	US-08-465-388-56
C 13	59.6	40.3	4164	1	US-08-204-675-1
C 14	59.6	40.3	4164	3	US-08-660-734-1
C 15	59.6	40.3	4164	4	US-08-796-364-1
C 16	59.6	40.3	4164	6	PCT-US95-02520-1
C 17	58.4	39.5	5178	3	US-08-474-169-2
C 18	56	37.8	88	1	US-08-144-602B-15
C 19	56	37.8	1023	1	US-08-198-446B-1
C 20	56	37.8	1023	3	US-08-870-693-1
C 21	56	37.8	2150	1	US-08-198-446B-10
C 22	56	37.8	2150	3	US-08-870-693-10
C 23	55	37.2	834	3	US-08-967-101-113
C 24	55	37.2	834	4	US-08-592-541-113
C 25	51.6	34.9	6505	3	US-08-793-610-5
C 26	51	34.5	3341	3	US-08-868-577-18

C 27	50	33.8	780	3	US-08-967-101-116	Sequence 116, App
C 28	50	33.8	780	4	US-08-592-541-116	Sequence 116, App
C 29	49.8	33.6	1200	1	US-08-011-398B-3	Sequence 3, Appl
C 30	49.8	33.6	1200	2	US-08-464-051-3	Sequence 3, Appl
C 31	49.8	33.6	1200	3	US-08-462-498-3	Sequence 3, Appl
C 32	49.8	33.6	6244	1	US-08-076-726-15	Sequence 15, Appl
C 33	49.8	33.6	6244	1	US-08-260-452-8	Sequence 8, Appl
C 34	49.8	33.6	6244	3	US-08-481-970-8	Sequence 8, Appl
C 35	49.8	33.6	6244	4	US-08-897-719-8	Sequence 8, Appl
C 36	48.2	32.6	2308	1	US-08-325-071-62	Sequence 62, Appl
C 37	48	32.4	1612	1	US-08-343-733A-1	Sequence 1, Appl
C 38	48	32.4	2764	4	US-08-485-971B-1	Sequence 1, Appl
C 39	47.2	31.9	2185	4	US-08-467-948A-3	Sequence 3, Appl
C 40	47	31.8	1249	2	US-08-463-115-35	Sequence 35, Appl
C 41	47	31.8	1249	2	US-08-465-388-35	Sequence 35, Appl
C 42	47	31.8	1997	4	US-08-667-809B-3	Sequence 3, Appl
C 43	47	31.8	2126	3	US-08-789-354-1	Sequence 1, Appl
C 44	47	31.8	2605	3	US-08-680-395-4	Sequence 4, Appl
C 45	47	31.8	5534	2	US-08-452-267-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-967-101-121/c
; Sequence 121, Application US/08967101
; Patent No. 5840540
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,101
; FILING DATE: 10-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-967-101-121

Query Match 46.2%; Score 68.4; DB 3; Length 793;
Best Local Similarity 87.2%; Pred. No. 3.8e-14;
Matches 75; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 48 GGTGGATCCCCGGGTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
Db 119 GCGGATCCCCGGGTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 60

QY 108 GGGGGGCTTAACCTAACTTTTGT 133
Db 59 GGGGGGCGGTACCACTTTTGT 34

RESULT 2
US-08-592-541-121/c
; Sequence 121, Application US/08592541
; Patent No. 5986054
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541

; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-592-541-121

Query Match 46.2%; Score 68.4; DB 4; Length 793;
Best Local Similarity 87.2%; Pred. No. 3.8e-14;
Matches 75; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 48 GGTGGATCCCCGGGTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
Db 119 GCGGATCCCCGGGTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 60

QY 108 GGGGGGCTTAACCTAACTTTTGT 133
Db 59 GGGGGGCGGTACCACTTTTGT 34

RESULT 3
US-08-992-334-1/c
; Sequence 1, Application US/08992334
; Patent No. 5919678
; GENERAL INFORMATION:
; APPLICANT: Gruss, Alexandra
; APPLICANT: Maguin, Emmanuelle
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE
; TITLE OF INVENTION: PLASMID

; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christie Parker & Hale, LLP
; STREET: 350 West Colorado Boulevard, Suite 500
; CITY: Pasadena
; STATE: California
; COUNTRY: United States
; ZIP: 91105

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992,334

; FILING DATE: 17-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,752
; FILING DATE: 24-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00248

; FILING DATE: 12-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/03034
; FILING DATE: 13-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Prout, D. Bruce
; REGISTRATION NUMBER: 20958
; REFERENCE/DOCKET NUMBER: C93:31779

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626) 795-9900
; TELEFAX: (626) 577-8800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3792 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
; IMMEDIATE SOURCE:
; CLONE: pg+host4
US-08-992-334-1

Query Match 45.5%; Score 67.4; DB 4; Length 3792;
Best Local Similarity 98.6%; Pred. No. 1.3e-13;
Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 47 TGGTGGATCCCCGGGTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
Db 3543 TAGTGGATCCCCGGGTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 3484

QY 107 GGGGGGGGCC 115
Db 3483 GGGGGGGGCC 3475

RESULT 4
US-08-302-752-1/c
; Sequence 1, Application US/08302752
; Patent No. 6025190
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: THERMOSENSIBLE PLASMID
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)


```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302.752
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9203034
; FILING DATE: 13-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR/93/00248
; FILING DATE: 12-MAR-1993
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3792 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-302-752-1

Query Match 45.5%; Score 67.4; DB 5; Length 3792;
Best Local Similarity 98.6%; Pred. NO. 1.3e-13;
Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 47 TGGTGGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
DB 3543 TAGTGGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 3484

QY 107 GGGGGGGCC 115
DB 3483 GGGGGGGCC 3475

RESULT 5
US-08-410-540-3
; Sequence 3, Application US/08410540
; Patent No. 5807678
; GENERAL INFORMATION:
; APPLICANT: Miller, Walter L.
; APPLICANT: Lin, Dong
; APPLICANT: Strauss III, Jerome F.
; TITLE OF INVENTION: IDENTIFICATION OF GENE MUTATIONS
; TITLE OF INVENTION: ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94306-2155
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/410.540
; FILING DATE: 23-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Neeley, Richard L.
; REGISTRATION NUMBER: 30,092
; REFERENCE/DOCKET NUMBER: UCAL-238/0005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 853 5070
; TELEFAX: 415 857 0663
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1098..1283
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1620..1733
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2047..2174
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2267..2425
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2567..2751
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2828..2921
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3031..3765
; FEATURE:
; NAME/KEY: -
; LOCATION: 1433..1434
; OTHER INFORMATION: /note= "interruption of sequence data"
; FEATURE:
; NAME/KEY: -
; LOCATION: 2208..2209
; OTHER INFORMATION: /note= "interruption of sequence data"
; FEATURE:
; NAME/KEY: -
; LOCATION: 2781..2782
; OTHER INFORMATION: /note= "interruption of sequence data"
; FEATURE:
; NAME/KEY: -
; LOCATION: 2956..2957
; OTHER INFORMATION: /note= "interruption of sequence data"
US-08-410-540-3

Query Match 45.5%; Score 67.4; DB 2; Length 4016;
Best Local Similarity 98.6%; Pred. NO. 1.3e-13;
Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 47 TGGTGGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
DB 3937 TAGTGGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 3996

QY 107 GGGGGGGCC 115
DB 3937 GGGGGGGCC 4005

RESULT 6
US-08-992-334-2/c
; Sequence 2, Application US/08992334
; Patent No. 5519678
; GENERAL INFORMATION:
; APPLICANT: Gruss, Alexandra
; APPLICANT: Maguin, Emmanuelle
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE
; TITLE OF INVENTION: PLASMID
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christie Parker & Hale, LLP
; STREET: 350 West Colorado Boulevard, Suite 500
; CITY: Pasadena
; STATE: California
; COUNTRY: United States

```

```
;
;
; ZIP: 91105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992,334
; FILING DATE: 17-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,752
; FILING DATE: 24-DEC-1994
; PRIOR APPLICATION NUMBER: PCT/FR93/00248
; APPLICATION DATA:
; FILING DATE: 12-MAR-1993
; APPLICATION NUMBER: FR FR92/03034
; FILING DATE: 13-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Prout, D. Bruce
; REGISTRATION NUMBER: 20958
; REFERENCE/DOCKET NUMBER: C93:31779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626) 795-9900
; TELEFAX: (626) 577-8800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5234 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-992-334-2

Query Match 45.5%; Score 67.4; DB 4; Length 5234;
Best Local Similarity 98.6%; Pred. No. 1.4e-13;
Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 47 TGGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
DB 4985 TAGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 4926

QY 107 GGGGGGGCC 115
DB 4925 GGGGGGGCC 4917

RESULT 7
US-08-302-752-2/c
; Sequence 2, Application US/08302752
; Patent No. 6025190
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: THERMOSENSIBLE PLASMID
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,752
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9203034
; FILING DATE: 13-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR/93/00248
; FILING DATE: 12-MAR-1993
; INFORMATION FOR SEQ ID NO: 2:
```

```
;
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5234 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-302-752-2

Query Match 45.5%; Score 67.4; DB 5; Length 5234;
Best Local Similarity 98.6%; Pred. No. 1.4e-13;
Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 47 TGGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
DB 4985 TAGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 4926

QY 107 GGGGGGGCC 115
DB 4925 GGGGGGGCC 4917

RESULT 8
US-08-992-334-3/c
; Sequence 3, Application US/08992334
; Patent No. 5919678
; GENERAL INFORMATION:
; APPLICANT: Gruss, Alexandra
; APPLICANT: Maguin, Emmanuelle
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE
; TITLE OF INVENTION: PLASMID
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESS: Christie Parker & Hale, LLP
; STREET: 350 West Colorado Boulevard, Suite 500
; CITY: Pasadena
; STATE: California
; COUNTRY: United States
; ZIP: 91105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992,334
; FILING DATE: 17-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,752
; FILING DATE: 24-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00248
; FILING DATE: 12-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR FR92/03034
; FILING DATE: 13-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Prout, D. Bruce
; REGISTRATION NUMBER: 20958
; REFERENCE/DOCKET NUMBER: C93:31779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626) 795-9900
; TELEFAX: (626) 577-8800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-992-334-3
```

```

; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
; US-08-793-610-6

Query Match 44.2%; Score 65.4; DB 3;
Best Local Similarity 98.5%; Pred. No. 7, 6e-13;
Matches 66; Conservative 0; Mismatches 1;

QY 49 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGAT
Db 5654 GGGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGAT
QY 109 GGGGGCC 115
Db 5714 GGGGGGCC 5720

RESULT 11
US-08-463-115-56
; Sequence 56, Application US/08463115
; Patent No. 5703221
; GENERAL INFORMATION:
; APPLICANT: WILLIAM JOHN MARTIN
; TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
; TITLE OF INVENTION: AND RELATED VACCINES
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IEM Compatible

```

```
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463.115
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/157,811
; FILING DATE: No. 5703221ember 23, 1993
; APPLICATION NUMBER: 07/887,502
; FILING DATE: May 22, 1992
; APPLICATION NUMBER: 07/704,814
; FILING DATE: May 23, 1991
; APPLICATION NUMBER: 07/763,039
; FILING DATE: September 20, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 213/301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 685 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION:
; US-08-463-115-56

Query Match 42.4%; Score 62.8; DB 2; Length 685;
Best Local Similarity 92.8%; Pred. No. 2.7e-12;
Matches 64; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 47 TGGTGGATCCCCGGCGTCCAGGATTCGATATCGATACCGTTCGACCTCGA 106
DB 148 TAGTGGATCCCCGGCGTCCAGGATTCGATATCGATACCGTTCGACCTCGA 207

QY 107 GGGGGGGGCC 115
DB 208 GGGGGGGGCC 216

RESULT 12
US-08-465-388-56
; Sequence 56, Application US/08465388
; Patent No. 5753488
; GENERAL INFORMATION:
; APPLICANT: WILLIAM JOHN MARTIN
; TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
; TITLE OF INVENTION: AND RELATED VACCINES
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LYON & LYON
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/465,388
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/157,811
; FILING DATE: No. 5753488ember 23, 1993
; APPLICATION NUMBER: 07/887,502
; FILING DATE: May 22, 1992
; APPLICATION NUMBER: 07/704,814
; FILING DATE: May 23, 1991
; APPLICATION NUMBER: 07/763,039
; FILING DATE: September 20, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 213/300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEX: (213) 955-0440
; TELEPHONE: 67-3510
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 685 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION:
; US-08-465-388-56

Query Match 42.4%; Score 62.8; DB 2; Length 685;
Best Local Similarity 92.8%; Pred. No. 2.7e-12;
Matches 64; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 47 TGGTGGATCCCCGGCGTCCAGGATTCGATATCGATACCGTTCGACCTCGA 106
DB 148 TAGTGGATCCCCGGCGTCCAGGATTCGATATCGATACCGTTCGACCTCGA 207

QY 107 GGGGGGGGCC 115
DB 208 GGGGGGGGCC 216

RESULT 13
US-08-204-675-1
; Sequence 1, Application US/08204675
; Patent No. 567170
; GENERAL INFORMATION:
; APPLICANT: Devine, Scott E.
; APPLICANT: Boeke, Jef D.
; APPLICANT: Braiterman, Lelita T.
; TITLE OF INVENTION: In Vitro Transposition of Artificial
; TITLE OF INVENTION: Transposons
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie, and Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,675
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 435
```

```
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kagan, Sarah A.
/ REGISTRATION NUMBER: 32,141
/ REFERENCE/DOCKET NUMBER: 01107.45501
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202.508.9100
/ TELEFAX: 202.508.9299
/ TELEX: 197430 BBMB UT
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4164 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ IMMEDIATE SOURCE:
/ CLONE: PAT-1
/ US-08-660-754-1

Query Match          40.3%; Score 59.6; DB 1; Length 4164;
Best Local Similarity 75.5%; Pred. No. 5.2e-11;
Matches 74; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 10 TTGAAAACCTAGATCTATTATTGACGCGCGCGCTCTAGAACTAGTGGATCCCGGGCTGCAGG 69
DB 2005 TCGAACATGTTACCGCGGCTTATCGATACCGTTCGACCTCGAG 2102

QY 70 AATTCGATATCAAGCTTATCGATACCGTTCGACCTCGAG 107
DB 2065 AATTCGATATCAAGCTTATCGATACCGTTCGACCTCGAG 2102

RESULT 15
US-08-796-364-1
/ Sequence 1, Application US/08796364
/ Patent No. 5968785
/ GENERAL INFORMATION:
/ APPLICANT: Devine, Scott E.
/ APPLICANT: Boeke, Jef D.
/ APPLICANT: Braiterman, Lelita T.
/ TITLE OF INVENTION: In Vitro Transposition of Artificial
/ NUMBER OF SEQUENCES: 7
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Banner, Birch, McKie, and Beckett
/ STREET: 1001 G Street, N.W.
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20001
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/796,364
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/204,675
/ FILING DATE: 02-MAR-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kagan, Sarah A.
/ REGISTRATION NUMBER: 32,141
/ REFERENCE/DOCKET NUMBER: 01107.45501
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202.508.9100
/ TELEFAX: 202.508.9299
/ TELEX: 197430 BBMB UT
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4164 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO

QY 10 TTGAAAACCTAGATCTATTATTGACGCGCGCGCTCTAGAACTAGTGGATCCCGGGCTGCAGG 69
DB 2005 TCGAACATGTTACCGCGGCTTATCGATACCGTTCGACCTCGAG 2102

QY 70 AATTCGATATCAAGCTTATCGATACCGTTCGACCTCGAG 107
DB 2065 AATTCGATATCAAGCTTATCGATACCGTTCGACCTCGAG 2102

RESULT 14
US-08-660-754-1
/ Sequence 1, Application US/08660754
/ Patent No. 5843772
/ GENERAL INFORMATION:
/ APPLICANT: Devine, Scott E.
/ APPLICANT: Boeke, Jef D.
/ APPLICANT: Braiterman, Lelita T.
/ TITLE OF INVENTION: In Vitro Transposition of Artificial
/ NUMBER OF SEQUENCES: 7
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Banner, Birch, McKie, and Beckett
/ STREET: 1001 G Street, N.W.
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20001
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/660,754
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/204,675
/ FILING DATE: 02-MAR-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kagan, Sarah A.
/ REGISTRATION NUMBER: 32,141
/ REFERENCE/DOCKET NUMBER: 01107.45501
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202.508.9100
/ TELEFAX: 202.508.9299
/ TELEX: 197430 BBMB UT
```

; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: PAT-1
US-08-796-364-1

Query Match 40.3%; Score 59.6; DB 4; Length 4164;
Best Local Similarity 75.5%; Pred. No. 5.2e-11;
Matches 74; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 10 TTGAAAACTAGATCTATTATTGCACGGCGGCCCATGGTGGATCCCGGGGCTGCAGG 69
Db 2005 TCGAACATGTTACCGGGTGGCGGCTCTAGAACTAGTGGATCCCGGGGCTGCAGG 2064
Qy 70 AATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
Db 2065 AATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102

Search completed: May 29, 2000, 22:09:05
Job time: 38746 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2000, 09:48:30 ; Search time 2276.24 Seconds
(without alignments)
290.102 Million cell updates/sec

Title: US-08-935-377-7
Perfect score: 148
Sequence: 1 GGCAAAATGTAAGAACTA.....TTGTTTTGTGGCCGGCC 148

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5142629 seqs, 2230885800 residues
Total number of hits satisfying chosen parameters: 10285240

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA:*

1:	/cgnl_6/ptodata/1/pna/PCTUS_COMB.seq.*
2:	/cgnl_6/ptodata/1/pna/US06_COMB.seq.*
3:	/cgnl_6/ptodata/1/pna/US07_COMB.seq.*
4:	/cgnl_6/ptodata/1/pna/US080_COMB.seq.*
5:	/cgnl_6/ptodata/1/pna/US081A_COMB.seq.*
6:	/cgnl_6/ptodata/1/pna/US081B_COMB.seq.*
7:	/cgnl_6/ptodata/1/pna/US081C_COMB.seq.*
8:	/cgnl_6/ptodata/1/pna/US081D_COMB.seq.*
9:	/cgnl_6/ptodata/1/pna/US082A_COMB.seq.*
10:	/cgnl_6/ptodata/1/pna/US082B_COMB.seq.*
11:	/cgnl_6/ptodata/1/pna/US082C_COMB.seq.*
12:	/cgnl_6/ptodata/1/pna/US083A_COMB.seq.*
13:	/cgnl_6/ptodata/1/pna/US083B_COMB.seq.*
14:	/cgnl_6/ptodata/1/pna/US084A_COMB.seq.*
15:	/cgnl_6/ptodata/1/pna/US084B_COMB.seq.*
16:	/cgnl_6/ptodata/1/pna/US084C_COMB.seq.*
17:	/cgnl_6/ptodata/1/pna/US085_COMB.seq.*
18:	/cgnl_6/ptodata/1/pna/US086_COMB.seq.*
19:	/cgnl_6/ptodata/1/pna/US087A_COMB.seq.*
20:	/cgnl_6/ptodata/1/pna/US087B_COMB.seq.*
21:	/cgnl_6/ptodata/1/pna/US087C_COMB.seq.*
22:	/cgnl_6/ptodata/1/pna/US088A_COMB.seq.*
23:	/cgnl_6/ptodata/1/pna/US088B_COMB.seq.*
24:	/cgnl_6/ptodata/1/pna/US088C_COMB.seq.*
25:	/cgnl_6/ptodata/1/pna/US088D_COMB.seq.*
26:	/cgnl_6/ptodata/1/pna/US089A_COMB.seq.*
27:	/cgnl_6/ptodata/1/pna/US089B_COMB.seq.*
28:	/cgnl_6/ptodata/1/pna/US089C_COMB.seq.*
29:	/cgnl_6/ptodata/1/pna/US089D_COMB.seq.*
30:	/cgnl_6/ptodata/1/pna/US089E_COMB.seq.*
31:	/cgnl_6/ptodata/1/pna/US089F_COMB.seq.*
32:	/cgnl_6/ptodata/1/pna/US089G_COMB.seq.*
33:	/cgnl_6/ptodata/1/pna/US089H_COMB.seq.*
34:	/cgnl_6/ptodata/1/pna/US089I_COMB.seq.*
35:	/cgnl_6/ptodata/1/pna/US090A_COMB.seq.*
36:	/cgnl_6/ptodata/1/pna/US090B_COMB.seq.*
37:	/cgnl_6/ptodata/1/pna/US090C_COMB.seq.*
38:	/cgnl_6/ptodata/1/pna/US090D_COMB.seq.*
39:	/cgnl_6/ptodata/1/pna/US090E_COMB.seq.*
40:	/cgnl_6/ptodata/1/pna/US091A_COMB.seq.*
41:	/cgnl_6/ptodata/1/pna/US091B_COMB.seq.*
42:	/cgnl_6/ptodata/1/pna/US091C_COMB.seq.*
43:	/cgnl_6/ptodata/1/pna/US092A_COMB.seq.*

44:	/cgnl_6/ptodata/1/pna/US092B_COMB.seq.*
45:	/cgnl_6/ptodata/1/pna/US092C_COMB.seq.*
46:	/cgnl_6/ptodata/1/pna/US092D_COMB.seq.*
47:	/cgnl_6/ptodata/1/pna/US092E_COMB.seq.*
48:	/cgnl_6/ptodata/1/pna/US093A_COMB.seq.*
49:	/cgnl_6/ptodata/1/pna/US093B_COMB.seq.*
50:	/cgnl_6/ptodata/1/pna/US093C_COMB.seq.*
51:	/cgnl_6/ptodata/1/pna/US093D_COMB.seq.*
52:	/cgnl_6/ptodata/1/pna/US093E_COMB.seq.*
53:	/cgnl_6/ptodata/1/pna/US093F_COMB.seq.*
54:	/cgnl_6/ptodata/1/pna/US094A_COMB.seq.*
55:	/cgnl_6/ptodata/1/pna/US094B_COMB.seq.*
56:	/cgnl_6/ptodata/1/pna/US094C_COMB.seq.*
57:	/cgnl_6/ptodata/1/pna/US094D_COMB.seq.*
58:	/cgnl_6/ptodata/1/pna/US094E_COMB.seq.*
59:	/cgnl_6/ptodata/1/pna/US094F_COMB.seq.*
60:	/cgnl_6/ptodata/1/pna/US094G_COMB.seq.*
61:	/cgnl_6/ptodata/1/pna/US094H_COMB.seq.*
62:	/cgnl_6/ptodata/1/pna/US094I_COMB.seq.*
63:	/cgnl_6/ptodata/1/pna/US094J_COMB.seq.*
64:	/cgnl_6/ptodata/1/pna/US094K_COMB.seq.*
65:	/cgnl_6/ptodata/1/pna/US094L_COMB.seq.*
66:	/cgnl_6/ptodata/1/pna/US094M_COMB.seq.*
67:	/cgnl_6/ptodata/1/pna/US094N_COMB.seq.*
68:	/cgnl_6/ptodata/1/pna/US094O_COMB.seq.*
69:	/cgnl_6/ptodata/1/pna/US094P_COMB.seq.*
70:	/cgnl_6/ptodata/1/pna/US094Q_COMB.seq.*
71:	/cgnl_6/ptodata/1/pna/US094R_COMB.seq.*
72:	/cgnl_6/ptodata/1/pna/US094S_COMB.seq.*
73:	/cgnl_6/ptodata/1/pna/US094T_COMB.seq.*
74:	/cgnl_6/ptodata/1/pna/US094U_COMB.seq.*
75:	/cgnl_6/ptodata/1/pna/US094V_COMB.seq.*
76:	/cgnl_6/ptodata/1/pna/US094W_COMB.seq.*
77:	/cgnl_6/ptodata/1/pna/US094X_COMB.seq.*
78:	/cgnl_6/ptodata/1/pna/US094Y_COMB.seq.*
79:	/cgnl_6/ptodata/1/pna/US094Z_COMB.seq.*
80:	/cgnl_6/ptodata/1/pna/US095A_COMB.seq.*
81:	/cgnl_6/ptodata/1/pna/US095B_COMB.seq.*
82:	/cgnl_6/ptodata/1/pna/US095C_COMB.seq.*
83:	/cgnl_6/ptodata/1/pna/US095D_COMB.seq.*
84:	/cgnl_6/ptodata/1/pna/US095E_COMB.seq.*
85:	/cgnl_6/ptodata/1/pna/US095F_COMB.seq.*
86:	/cgnl_6/ptodata/1/pna/US095G_COMB.seq.*
87:	/cgnl_6/ptodata/1/pna/US095H_COMB.seq.*
88:	/cgnl_6/ptodata/1/pna/US095I_COMB.seq.*
89:	/cgnl_6/ptodata/1/pna/US095J_COMB.seq.*
90:	/cgnl_6/ptodata/1/pna/US095K_COMB.seq.*
91:	/cgnl_6/ptodata/1/pna/US095L_COMB.seq.*
92:	/cgnl_6/ptodata/1/pna/US095M_COMB.seq.*
93:	/cgnl_6/ptodata/1/pna/US095N_COMB.seq.*
94:	/cgnl_6/ptodata/1/pna/US095O_COMB.seq.*
95:	/cgnl_6/ptodata/1/pna/US095P_COMB.seq.*
96:	/cgnl_6/ptodata/1/pna/US095Q_COMB.seq.*
97:	/cgnl_6/ptodata/1/pna/US095R_COMB.seq.*
98:	/cgnl_6/ptodata/1/pna/US095S_COMB.seq.*
99:	/cgnl_6/ptodata/1/pna/US095T_COMB.seq.*
100:	/cgnl_6/ptodata/1/pna/US095U_COMB.seq.*
101:	/cgnl_6/ptodata/1/pna/US095V_COMB.seq.*
102:	/cgnl_6/ptodata/1/pna/US095W_COMB.seq.*
103:	/cgnl_6/ptodata/1/pna/US095X_COMB.seq.*
104:	/cgnl_6/ptodata/1/pna/US095Y_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	148	100.0	148	28	US-08-935-377-7
2	137	92.6	149	28	US-08-935-377-8

Sequence 7, Appli
Sequence 8, Appli

ALIGNMENTS

```

1
RESULT
US-08-935-377-7
; Sequence 7, Application US/08935377
;
; GENERAL INFORMATION:
;
; APPLICANT: Zauderer, Maurice
;
; TITLE OF INVENTION: T Cells Specific for Target Antigens and
;
; TITLE OF INVENTION: Vaccines Based Thereon
;
; NUMBER OF SEQUENCES: 37
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
;
; STREET: 1100 New York Avenue, N.W., Suite 500
;
; CITY: Washington
;
; STATE: D. C.
;
; COUNTRY: USA
;
; ZIP: 20005
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/935,377
;
; FILING DATE: 22-SEP-1997
;
; CLASSIFICATION: 424
;
; ATTORNEY/AGENT INFORMATION:

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/935,377
FILING DATE: 22-SEP-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 149 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA

```


Query Match 92.6%; Score 137; DB 28; Length 149;
Best Local Similarity 99.3%; Pred. No. 5e-36;
Matches 148; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GCCCAAAATTCAGAACTAGATCTATTATTGCACGGCGGCCCATG-GTGGATCCCC 59
DB 1 GCCCAAAATTCAGAACTAGATCTATTATTGCACGGCGGCCCATGAGTGGATCCCC 60
QY 60 GGGCTGCAGGAATTCGATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCGCTAAC 119
DB 61 GGGCTGCAGGAATTCGATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCGCTAAC 120
QY 120 TAACTAAATTTGTTTTGTGGCCCGGCC 148
DB 121 TAACTAAATTTGTTTTGTGGCCCGGCC 149

RESULT 3
US-08-935-377-9
; Sequence 9, Application US/08935377
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; TITLE OF INVENTION: T Cells Specific for Target Antigens and
; TITLE OF INVENTION: Vaccines Based Thereon
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,377
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-935-377-9

Query Match 91.9%; Score 136; DB 28; Length 150;
Best Local Similarity 98.7%; Pred. No. 1.1e-35;
Matches 148; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 GCCCAAAATTCAGAACTAGATCTATTATTGCACGGCGGCCCATG-GTGGATCCCC 58
DB 1 GCCCAAAATTCAGAACTAGATCTATTATTGCACGGCGGCCCATGACGTGGATCCCC 60
QY 59 CGGGCTGCAGGAATTCGATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCGCTAA 118
DB 61 CGGGCTGCAGGAATTCGATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCGCTAA 120
QY 119 CTAACAAATTTGTTTTGTGGCCCGGCC 148
DB 119 CTAACAAATTTGTTTTGTGGCCCGGCC 149

DB 121 CTAACAAATTTGTTTTGTGGCCCGGCC 150

RESULT 4
US-08-935-377-6
; Sequence 6, Application US/08935377
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; TITLE OF INVENTION: T Cells Specific for Target Antigens and
; TITLE OF INVENTION: Vaccines Based Thereon
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,377
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-935-377-6

Query Match 89.2%; Score 132; DB 28; Length 145;
Best Local Similarity 98.0%; Pred. No. 2.4e-34;
Matches 145; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 GCCCAAAATTCAGAACTAGATCTATTATTGCACGGCGGCCCATGTTGGATCCCCCG 60
DB 1 GCCCAAAATTCAGAACTAGATCTATTATTGCACGGCGGCCCATGTTGGATCCCCCG 57
QY 61 GGGCTGCAGGAATTCGATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCGCTAACT 120
DB 58 GGGCTGCAGGAATTCGATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCGCTAACT 117
QY 121 AACTAAATTTGTTTTGTGGCCCGGCC 148
DB 118 AACTAAATTTGTTTTGTGGCCCGGCC 145

RESULT 5
US-09-123-912-91
; Sequence 91, Application US/09123912A
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C1
; CURRENT APPLICATION NUMBER: US/09/123,912A
; CURRENT FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: 09/040,802
; PRIOR FILING DATE: 1998-03-18

Query Match 47.7% Score 70.6; DB 40; Length 858;
Best Local Similarity 84.9% Pred. No. 1.6e-13;
Matches 79; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 23 TCTATTTTTCACGCGCGCCCATGTGGATCCCCGGGCTCGAGGAATTCGATCA 82

Qy 23 TCTATTATTGACGGCGGCCATGTGGTATCCCGGCTGCAGGAATTCGATCAA 82
 ||| ||| ||| | | ||||| ||||| ||||| ||||| ||||| |||||
Db 331 tcaacgaattccaccacactgactagtgtatccccggctgcaggaaattcgatatcaa 390

Qy 83 GCTTATCGATACCCTGACCTCGAGGGGGGGCC 115

```

RESULT          9
US-08-693-573-4
; Sequence 4, Application US/08693573
; GENERAL INFORMATION:
; APPLICANT: HOSOI, Shigeru,
; APPLICANT: FUKAMI, Tadashi,
; APPLICANT: KOJIMA, Makiko
; TITLE OF INVENTION: Method of Determining Base Sequence of
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARY CUSHMAN
; ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
;

```

STREET: Ninth Floor, 1100 New York Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 720Kb storage
COMPUTER: IBM PC/XT/AT compatibles
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1 or ASCII editors
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/693,573
FILING DATE: 07-AUG-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/377,383
FILING DATE: 25-JAN-1995
APPLICATION NUMBER: JP6971/1994
FILING DATE: 26-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Thomas G. Wiseman
REGISTRATION NUMBER: 35046
TELEPHONE: (202)861-3000
TELEFAX: (202)822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid Synthetic DNA
US-08-693-573-4

Query Match 46.9%; Score 69.4; DB 18; Length 224;
Best Local Similarity 87.4%; Pred. No. 2.9e-13;
Matches 76; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 47 TGGTGGATCCCCGGCGTCGAGGAATTCGATCAAGCTTATCGATACCGTCGACCTCGA 106
DB 86 TAGTGGATCCCCGGCGTCGAGGAATTCGATCAAGCTTATCGATACCGTCGACCTCGA 145

QY 107 GGGGGGGCTAACTAACTAATTTGTT 133
DB 146 GGGGGGGCGGTACCCAGCTTTGTT 172

RESULT 10
US-08-431-048A-121/c.
Sequence 121, Application US/08431048A
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 132
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patent Administrator, TESTA, HURWITZ &
ADDRESSEE: THIBEAULT, LLP
STREET: 125 High Street
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/431,048A
FILING DATE: 28-APR-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Twomey, Michael J
REGISTRATION NUMBER: 38,349
REFERENCE/DOCKET NUMBER: CAN-002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 793 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-431-048A-121

Query Match 46.2%; Score 68.4; DB 15; Length 793;
Best Local Similarity 87.2%; Pred. No. 8.8e-13;
Matches 75; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 48 GGTGGATCCCCGGCGTCGAGGAATTCGATCAAGCTTATCGATACCGTCGACCTCGAG 107
DB 119 GCGGGATCCCCGGCGTCGAGGAATTCGATCAAGCTTATCGATACCGTCGACCTCGAG 60

QY 108 GGGGGCGCTAACTAACTAATTTGTT 133
DB 59 GGGGGCGCGGTACCCAGCTTTGTT 34

RESULT 11
US-08-496-841-121/c
Sequence 121, Application US/08496841
GENERAL INFORMATION:
APPLICANT: St. George-Hyslop, Peter
APPLICANT: ROMMENS, JOHANNA M.
APPLICANT: FRASER, PAUL E.
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 160
CORRESPONDENCE ADDRESS:
ADDRESSEE: Testa, Hurwitz & Thibault
STREET: High Street tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/496,841
FILING DATE: 28-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Twomey, Michael J
REGISTRATION NUMBER: 38349
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7362
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 793 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-496-841-121


```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,523
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-124-523-121

Query Match 46.2%; Score 68.4; DB 40; Length 793;
Best Local Similarity 87.2%; Pred. No. 8.8e-13;
Matches 75; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 48 GGTGGATCCCCGGGCTCGAGATTGATCAAGCTTATCGATACCGTCGACCTCGAG 107
   |||||||
Db 119 GCGGATCCCCGGGCTCGAGATTGATCAAGCTTATCGATACCGTCGACCTCGAG 60
   |||||||

QY 108 GGGGGCCCTAACTAACTAATTTGTT 133
   |||||||
Db 59 GGGGGCCCGGTACCGACTTTGTT 34
   |||||||

RESULT 15
US-09-124-698-121/c
; Sequence 121, Application US/09124698
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMWENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSES: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,698
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 base pairs

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:13:15 ; Search time 2192.43 Seconds
(without alignments)
273.613 Million cell updates/sec

Title: US-08-935-377-7
Perfect score: 148
Sequence: 1 GGCACAAATGAAAACTA.....TTGTTTTGTGGCCCGGCC 148

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues 9714632
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: em_est20:*
50: em_est21:*
51: em_est22:*
52: gb_est30:*
53: gb_est31:*
54: gb_est32:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: em_gss5:*
93: em_gss6:*
94: gb_gss7:*
95: gb_gss8:*
96: gb_gss9:*
97: em_gss7:*
98: em_gss8:*
99: gb_gss11:*
100: gb_gss10:*
101: em_gss9:*
102: em_gss10:*
103: em_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES source
Location/Qualifiers
1..682
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2301123"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT 156 a 187 c 187 g 151 t 1 others
ORIGIN

Query Match 47.6%; Score 70.4; DB 96; Length 682;
Best Local Similarity 92.5%; Fred. No. 2e-13; 6; Indels 0; Gaps 0;
Matches 74; Conservative 0; Mismatches 0;

QY 47 TGTGTGATCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGA 106
|||||
Db 80 TAGTGTATCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGA 21
|||||
QY 107 GGGGGGGCCCTAACTACTAA 126
|||||
Db 20 GGGGGGGCCCGAGTACCCAA 1
|||||

RESULT 2
AQ009167/c DNA GSS 27-JUN-1998
LOCUS
DEFINITION
CIT-HSP-2281N13.TF CIT-HSP Homo sapiens genomic clone 2281N13,
genomic survey sequence.
ACCESSION AQ009167
VERSION AQ009167.1 GI:3128599
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 608)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J.,
Shizuya,H., Simon,M., and Venter,J.C.
Use of a human BAC End Sequence Database for Sequence-Ready Map
Building
Unpublished (1997)
Other GSSs: CIT-HSP-2281N13.TR.1 CIT-HSP-2281N13.TF
CIT-HSP-2281N13.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.
FEATURES source
Location/Qualifiers
1..608

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES source
Location/Qualifiers
1..682
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2301123"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT 156 a 187 c 187 g 151 t 1 others
ORIGIN

Query Match 47.6%; Score 70.4; DB 96; Length 682;
Best Local Similarity 92.5%; Fred. No. 2e-13; 6; Indels 0; Gaps 0;
Matches 74; Conservative 0; Mismatches 0;

QY 47 TGTGTGATCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGA 106
|||||
Db 80 TAGTGTATCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGA 21
|||||
QY 107 GGGGGGGCCCTAACTACTAA 126
|||||
Db 20 GGGGGGGCCCGAGTACCCAA 1
|||||

RESULT 2
AQ009167/c DNA GSS 20-AUG-1998
LOCUS
DEFINITION
CIT-HSP-2301L23.TF CIT-HSP Homo sapiens genomic clone 2301L23,
genomic survey sequence.
ACCESSION AQ074693
VERSION AQ074693.1 GI:3436811
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 682)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M., and
Venter,J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)
Contact: Mark Adams

ALIGNMENTS
AQ074693 682 bp DNA GSS
CIT-HSP-2301L23.TF CIT-HSP Homo sapiens genomic clone 2301L23,
genomic survey sequence.
AQ074693.1 GI:3436811
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 682)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M., and
Venter,J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)
Contact: Mark Adams

Best Local Similarity 97.2%; Pred. No. 1.4e-12; Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	47	TGTTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA	106						
Db	77	TAGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA	18						
QY	107	GGGGGGGGCCTA	117						
Db	17	GGGGGGGGCCCA	7						
RESULT 4									
B54292 108 bp DNA GSS 20-JUN-1998									
CIT-HSP-2017M1.TR CIT-HSP Homo sapiens genomic clone 2017M1, genomic survey sequence.									
B54292									
B54292.1 GI:2608626									
GSS.									
human.									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Mammalia;									
Eutheria; Primates; Catarrhini; Hominidae; Homo.									
1 (bases 1 to 108)									
Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,									
Golden,K., Berry,K., Granger,D., Suh,E., Wibie,C., Shizuya,H.,									
Simon,M. and Venter,J.C.									
Use of a random BAC End Sequence Database for Sequence-Ready Map									
Building									
Unpublished (1997)									
Contact: Mark Adams									
Department of Eukaryotic Genomics									
The Institute for Genomic Research									
9712 Medical Center Dr., Rockville, MD 20850, USA									
Tel: 301 838 0200									
Fax: 301 838 0208									
Email: mdadams@tigr.org									
Clones are available from Research Genetics (info@resgen.com). BAC									
end search page:									
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html									
Seq primer: M13 Reverse									
Class: BAC ends.									
Location/Qualifiers									
1..108									
/organism="Homo sapiens"									
/db_xref="GDB:7044003"									
/db_xref="taxon:9606"									
/clone="2017M1"									
/clone_lib="CIT-HSP"									
/sex="Male"									
/cell_type="Sperm"									
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:									
HindIII"									
20 a 35 c 34 g 19 t									
BASE COUNT									
ORIGIN									
Query Match 45.5%; Score 67.4; DB 84; Length 108;									
Best Local Similarity 98.8%; Pred. No. 1.5e-12; Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	47	TGTTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA	106						
Db	32	TAGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA	91						
QY	107	GGGGGGGGCC	i15						
Db	92	GGGGGGGGCC	100						
RESULT 5									
AU037147									

Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13-21;
Seq: BAC ends.

Location/Qualifiers
1. .141

```
1. 141
/organism="Homo sapiens"
/db_xref="GDB:710398"
/db_xref="taxon:9606"
/clone="217N2"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
```

```

/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
HindIII"

```

HAWAII		
BASE COUNT	24 a	41 c
ORIGIN	24 a	41 g
		35 t

```

Query Match          45.5%; Score 67.4; DB 96; Length 141;
Best Local Similarity 98.6%; Pred. No. 1.6e-12;
Mismatched          0; Mismatches 1; Indels 0; Gaps 0;
Conservative

```

QY 47 TGGTGGATCCCCCGGCTGCAGGAATTGATATCAAGTTATCGATACCGTCGACCTCGA 106
- - - - -
25 TATGCTTGCCGCCCGCCTCCGCAATTCGATATCAAGTTATCGATACCGTCGACCTCGA 16

107 GGGGGGCC 115

QY	107	GGGGGGGGCC
Dp	15	GGGGGGGGCC

RESULT 7

NEGOTI
B80294

LOCUS
DEFINITION

100

ACCESSION
NUMBER

VERSION
KEYWORDS

KEYWORDS
SOURCE

ORGANISM

REFERENCE

REFERENCE AUTHORS

222

TITLE

JOURNAL

COMMENT	COUNT
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	1
10	1
11	1
12	1
13	1
14	1
15	1
16	1
17	1
18	1
19	1
20	1
21	1
22	1
23	1
24	1
25	1
26	1
27	1
28	1
29	1
30	1
31	1
32	1
33	1
34	1
35	1
36	1
37	1
38	1
39	1
40	1
41	1
42	1
43	1
44	1
45	1
46	1
47	1
48	1
49	1
50	1
51	1
52	1
53	1
54	1
55	1
56	1
57	1
58	1
59	1
60	1
61	1
62	1
63	1
64	1
65	1
66	1
67	1
68	1
69	1
70	1
71	1
72	1
73	1
74	1
75	1
76	1
77	1
78	1
79	1
80	1
81	1
82	1
83	1
84	1
85	1
86	1
87	1
88	1
89	1
90	1
91	1
92	1
93	1
94	1
95	1
96	1
97	1
98	1
99	1
100	1

SECRET

FEATURES SOURCE

0450

100

QY	47	TGTTGGATCCCGGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA	108
I			
Db	101	TAGTGGATCCCGGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA	42
QY	107	GGGGGGGCC	115
Db	41	GGGGGGGCC	33
RESULT	14		
B65534/C			
LOCUS	B65534	519 bp	DNA
DEFINITION	Cit-HSP-2022D24.1f Cit-HSP Homo sapiens genomic clone 2022D24.1	GSS	21-JUN-1998
ACCESSION	B65534		
VERSION	B65534		
KEYWORDS	GSS		
SOURCE	B65534.1	GI:2639512	
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	1 (bases 1 to 519)		
	Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.		
TITLE	Use of a random BAC End Sequence Database for Sequence-Ready M Building		
JOURNAL	Unpublished (1997)		
COMMENT	Other_GSSs: CIT-HSP-2022D24.TR Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: mdadams@tigr.org Clones are available from Research Genetics (info@resgen.com). end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.h Seq primer: Mi3-21 Class: BAC ends.		
FEATURES	Location/Qualifiers		
source	1..519		
	/organism="Homo sapiens"		
	/db_xref="GDB:7045730"		
	/db_xref="taxon:9606"		
	/clone="2022D24"		
	/clone_lib="CIT-HSP"		
	/sex="Male"		
	/cell_type="Sperm"		
	/note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII"		
BASE COUNT	110 a 137 c 148 g 124 t		
ORIGIN			
	Query Match 45.5%; Score 67.4; DB 84; Length 519;		
	Best Local Similarity 98.6%; Pred. No. 2e-12;		
	Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps		
QY	47	TGTTGGATCCCGGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA	106
I			
Db	84	TAGTGGATCCCGGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA	25
QY	107	GGGGGGGCC	115
Db	24	GGGGGGGCC	16
RESULT	15		
AQ012629			
LOCUS	AQ012629	547 bp	DNA
		GSS	06-JUN-1998

```

Query Match      45.5%; Score 67.4; DB 96; Length 547;
Best Local Similarity 98.6%; Pred. No. 2e-12;
Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 47 TGGTGGATCCCCGGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 106
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 73 TAGTGGATCCCCGGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 132

QY 107 GGGGGGGGCC 115
      | | | | | | | |
Db 133 GGGGGGGGCC 141

Search completed: May 29, 2000, 21:13:16
Job time: 36515 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:35:34 ; Search time 1214.87 Seconds
(without alignments)
934.219 Million cell updates/sec

Title: US-08-935-377-8
Perfect score: 149
Sequence: 1 GCCCAAAATTGAAAACTA.....TTGTTTTTGTGGCCCGGCC 149

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, 3808571567 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_om.*
4: gb_ov.*
5: gb_pat.*
6: gb_ph.*
7: gb_pl1.*
8: gb_pl2.*
9: gb_pr1.*
10: gb_pr2.*
11: gb_pr3.*
12: gb_ro.*
13: gb_sts.*
14: gb_sy.*
15: gb_un.*
16: gb_vl.*
17: em_fun.*
18: em_humi.*
19: em_hum2.*
20: em_in.*
21: em_om.*
22: em_or.*
23: em_ov.*
24: em_pat.*
25: em_ph.*
26: em_pl.*
27: em_ro.*
28: em_sts.*
29: em_sy.*
30: em_un.*
31: em_vl.*
32: gb_htg1.*
33: gb_htg2.*
34: gb_in1.*
35: gb_in2.*
36: em_ba1.*
37: em_ba2.*
38: em_hum3.*
39: em_hum4.*
40: gb_pr4.*
41: gb_htg3.*
42: gb_htg4.*
43: gb_htg5.*

44: gb_htg6.*
45: gb_htg7.*
46: em_htg1.*
47: em_htg2.*
48: em_htg3.*
49: em_hum5.*
50: gb_pl3.*
51: gb_pr5.*
52: gb_htg8.*
53: gb_htg9.*
54: gb_htg10.*
55: gb_htg11.*
56: gb_htg12.*
57: gb_htg13.*
58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	73.6	49.4	5639	14	CVU14122
C 2	71.2	47.8	259	5	A44281
C 3	70	47.0	2958	14	ARBLKSM
C 4	70	47.0	2958	14	ARBLKSP
C 5	70	47.0	2961	14	ARBL2KSM
C 6	70	47.0	2961	14	ARBL2KSP
C 7	70	47.0	2964	14	SYNBLKSMV
C 8	70	47.0	2964	14	SYNBLKSPV
C 9	70	47.0	3306	14	SYNBLKSPV
C 10	70	47.0	4144	14	XXU35131
C 11	70	47.0	4267	14	PRS304
C 12	70	47.0	4289	14	XXU35136
C 13	70	47.0	4373	14	PRS306
C 14	70	47.0	4443	14	PRS303
C 15	70	47.0	4670	14	ASAJ5326
C 16	70	47.0	4670	14	ASAJ5329
C 17	70	47.0	4707	14	XXU02374
C 18	70	47.0	4768	14	XXU25061
C 19	70	47.0	4783	14	PRS314
C 20	70	47.0	4887	14	PRS316
C 21	70	47.0	4950	14	XXU25060
C 22	70	47.0	4967	14	PRS313
C 23	70	47.0	5144	14	CVU23751
C 24	70	47.0	5187	14	U34887
C 25	70	47.0	5228	14	XXU25059
C 26	70	47.0	5504	14	PRS305
C 27	70	47.0	5634	14	CVU14125
C 28	70	47.0	6018	14	PRS315
C 29	70	47.0	6340	14	ASAJ5323
C 30	70	47.0	9655	14	SYNPR8V
C 31	68.8	46.2	3240	8	AF015771
C 32	68	45.6	144	14	SYNPLKRA
C 33	68	45.6	147	14	SYNPLKRB
C 34	68	45.6	2958	14	ARBLKSM
C 35	68	45.6	2958	14	ARBLSKP
C 36	68	45.6	2961	14	ARBL2SKM
C 37	68	45.6	2961	14	ARBL2SRP
C 38	68	45.6	2964	14	SYNBLDRPV
C 39	68	45.6	2964	14	SYNBLSKMV
C 40	68	45.6	3062	14	CVU47947
C 41	68	45.6	3228	14	CVU7829
C 42	68	45.6	3345	14	XXU35235
C 43	68	45.6	3357	14	XXU35132
C 44	68	45.6	3357	14	XXU35132
C 45	68	45.6	3574	14	AF139061

ALIGNMENTS


```

TITLE      pBluescript II: gene mapping vectors
JOURNAL    Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE    90067967
FEATURES   Location/Qualifiers
            source          1..2958
                        /organism="synthetic construct"
                        /db_xref="taxon:32630"
            misc_feature    1..2958
                        /note="phagemid pBluescript KS(-)"
BASE COUNT 708 a 754 c 731 g 765 t
ORIGIN

Query Match      47.0%; Score 70; DB 14; Length 2958;
Best Local Similarity 88.4%; Pred. No. 8e-12;
Matches 76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
      |||||||
DB 686 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 745
      |||||||
QY 109 GGGGGCCCTAACTAACTAATTTGTT 134
      |||||||
DB 746 GGGGGCCCGGTACCCAGCTTTTGT 771
      |||||||

RESULT 4
ARBLKSP
LOCUS      ARBLKSP      2958 bp      DNA      circular      SYN      11-JUN-1998
DEFINITION pBluescript KS(+) vector DNA, phagemid excised from lambda ZAP.
ACCESSION X52331
VERSION   X52331.1 GI:58065
KEYWORDS  artificial sequence; cloning vector; expression vector; vector.
SOURCE    synthetic construct.
ORGANISM  artificial sequence.
REFERENCE 1 (bases 1 to 2958)
AUTHORS   Thomas,E.A.
TITLE     Direct Submission
JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
          11099 North Torrey Pines Rd., La Jolla, CA 92037, USA
REFERENCE 2 (bases 1 to 2958)
AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
          excision properties
JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE   88319944
REFERENCE 3 (bases 1 to 2958)
AUTHORS   Altling-Wees,M.A. and Short,J.M.
TITLE     pBluescript II: gene mapping vectors
JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE   90067967
ORGANISM  artificial sequence.
REFERENCE 1 (bases 1 to 2958)
AUTHORS   Thomas,E.A.
TITLE     Direct Submission
JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
          11099 North Torrey Pines Rd., La Jolla, CA 92037, USA
REFERENCE 2 (bases 1 to 2958)
AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
          excision properties
JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE   88319944
REFERENCE 3 (bases 1 to 2958)
AUTHORS   Altling-Wees,M.A. and Short,J.M.
TITLE     pBluescript II: gene mapping vectors
JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE   90067967
REFERENCE 4 (bases 1 to 2958)
AUTHORS   Lampe,D.J., Grant,T.E. and Robertson,H.M.
TITLE     Factors affecting transposition of the Himar1 mariner transposon in
          vitro
JOURNAL   Genetics 149 (1), 179-187 (1998)
MEDLINE   98250682
FEATURES   Location/Qualifiers
            source          1..2958
                        /organism="synthetic construct"
                        /db_xref="taxon:32630"
            misc_feature    1..2958
                        /note="phagemid pBluescript KS(+)"
BASE COUNT 749 a 734 c 751 g 724 t
ORIGIN

Query Match      47.0%; Score 70; DB 14; Length 2958;
Best Local Similarity 88.4%; Pred. No. 8e-12;
Matches 76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
      |||||||

```

```

      |||||||
DB 686 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 745
      |||||||
QY 109 GGGGGCCCTAACTAACTAATTTGTT 134
      |||||||
DB 746 GGGGGCCCGGTACCCAGCTTTTGT 771
      |||||||

RESULT 5
ARBL2KSP
LOCUS      ARBL2KSP      2961 bp      DNA      circular      SYN      10-MAY-1995
DEFINITION pBluescript II KS(-) vector DNA, phagemid excised from lambda
          ZAPII.
ACCESSION X52329
VERSION   X52329.1 GI:58060
KEYWORDS  artificial sequence; cloning vector; expression vector; vector.
SOURCE    synthetic construct.
ORGANISM  artificial sequence.
REFERENCE 1 (bases 1 to 2961)
AUTHORS   Thomas,E.A.
TITLE     Direct Submission
JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
          11099 North Torrey Pines Rd., La Jolla, CA 92037, USA
REFERENCE 2 (bases 1 to 2961)
AUTHORS   Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
TITLE     Lambda ZAP: a bacteriophage lambda expression vector with in vivo
          excision properties
JOURNAL   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE   88319944
REFERENCE 3 (bases 1 to 2961)
AUTHORS   Altling-Wees,M.A. and Short,J.M.
TITLE     pBluescript II: gene mapping vectors
JOURNAL   Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE   90067967
FEATURES   Location/Qualifiers
            source          1..2961
                        /organism="synthetic construct"
                        /db_xref="taxon:32630"
            misc_feature    1..2961
                        /note="phagemid pBluescriptII KS(-)"
BASE COUNT 706 a 758 c 735 g 762 t
ORIGIN

Query Match      47.0%; Score 70; DB 14; Length 2961;
Best Local Similarity 88.4%; Pred. No. 8e-12;
Matches 76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
      |||||||
DB 686 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 745
      |||||||
QY 109 GGGGGCCCTAACTAACTAATTTGTT 134
      |||||||
DB 746 GGGGGCCCGGTACCCAGCTTTTGT 771
      |||||||

RESULT 6
ARBL2KSP
LOCUS      ARBL2KSP      2961 bp      DNA      circular      SYN      10-MAY-1995
DEFINITION pBluescript II KS(+) vector DNA, phagemid excised from lambda
          ZAPII.
ACCESSION X52327
VERSION   X52327.1 GI:58061
KEYWORDS  artificial sequence; cloning vector; expression vector; vector.
SOURCE    synthetic construct.
ORGANISM  artificial sequence.
REFERENCE 1 (bases 1 to 2961)
AUTHORS   Thomas,E.A.
TITLE     Direct Submission
JOURNAL   Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,

```



```

VERSION      L08785.1  GI:310729
KEYWORDS     Synthetic construct DNA.
SOURCE       Synthetic construct
ORGANISM     artificial sequence.
REFERENCE    1 (bases 1 to 2964)
AUTHORS      Gilbert,W.
JOURNAL      Obtained from VecBase 3.0
COMMENT      Unpublished (1991)
              These data and their annotation were supplied to GenBank by Will
              Gilbert under the auspices of the GenBank Curator Program.
ENTRY       Bluescribe KS Plus - Cloning vector
TITLE       Bluescribe KS Plus - Cloning vector
DATE        28-JAN-1987
#sequence   02-FEB-1987
#sequence   04-MAR-1987
#sequence   03-APR-1987
ACCESSION   V00078
SOURCE      artificial
REFERENCE    #number 1
              #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
              #Journal Gene (1987) in press
              #citation Sequence data from Stratagene
              #comment sequence correction according to Stratagene COMMENT
              Obtained from Stratagene on floppy disc.
              Revised 02-FEB-1987 by F. Pfeiffer.
              1409/10 'AT' to 'TA' to match revised sequence of pBR322
              .Revised 4-MAR-1987 to match sequence of pUC19 on request
              of Stratagene
              Polylinker region revised 03-APR-1987 according to Stratagene
              COMMENT
              The stand shown corresponds to pUC19c.
              As in the published sequence of pUC19c, The M13mp19 lacZ region
              is on the complementary strand.
              COMMENT
              This vector contains the fl origin so that the plus strand
              can be obtained upon fl superinfection.
              KEYWORDS
              CROSSREFERENCE
              #parent
              VecBase(3):BlueM13p
              #parent
              VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
              VecSource(3):bgalks, GenBank(50):PF1
              #brother
              VecBase(3):BlueKSm, VecBase(3):BlueSKp
PARENT
  Features of BlueKSp (2964 bp)
    residue source
    3- 458 5943-5488 (c) phage f1
    460- 824 236- 400 pUC19c
    626- 645 1- 20 T7 promoter
    653- 760 108- 1 (c) BlueKS-polylinker
    772- 791 20- 1 (c) T3 promoter
    795-2964 448-2617 pUC19c
    Conflict (cfl) and Mutations (mut): none
PARENT
  Features of BlueKSp (2964 bp)
    residue source
    3- 458 5943-5488 (c) phage f1
    460- 824 449- 285 (c) pUC19
    626- 645 1- 20 T7 promoter
    653- 760 108- 1 (c) BlueKS-polylinker
    772- 791 20- 1 (c) T3 promoter
    795-1031 237- 1 (c) pUC19
    1032-2964 2686- 754 (c) pUC19
    Conflict (cfl) and Mutations (mut): none
FEATURE      643 start of T7-RNA synthesis
              774 (c) start of T3-RNA synthesis

1976-2764 789-1 (c) Ap-R; b-lactamase
POLYLINKER
  KpnI-DraII-ApaI-XhoI-SalI-ClaI-HindIII-EcoRV-EcoRI-PstI-
  SmaI-BamHI-SpeI-XbaI-NotI-XmaIII-BstXI-SacII-SacI SELECTION
  #resistance Ap
  #indicator beta-galactosidase
SUMMARY      BlueKSp #length 2964 #checksum 690.
              Location/Qualifiers
              1..2964
              /organism="synthetic construct"
              /db_xref="taxon:32630"
BASE COUNT   750 a 736 c 755 g 723 t
ORIGIN
Query Match 47.0%; Score 70; DB 14; Length 2964;
Best Local Similarity 88.4%; Pred. No. 8e-12;
Matches 76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 49 AGTGGATCCCCGGCGTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
Db 686 AGTGGATCCCCGGCGTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 745
QY 109 GGGGGCCCTACTACTAATTTTGT 134
Db 746 GGGGGCCCGGTACCCAGCTTTTGT 771
RESULT 9
SYNBPEN66/c
LOCUS        SYNBPEN66 3306 bp DNA circular SYN 06-FEB-1999
DEFINITION   Cloning vector pBEN66 DNA for aminoglycoside 3'-phosphotransferase,
              beta-lactamase, complete cds.
ACCESSION    D85525
VERSION      D85525.1 GI:1345433
KEYWORDS     plasmid; aminoglycoside 3'-phosphotransferase; beta-lactamase.
SOURCE       Cloning vector pBEN66 (lab_host:E.coli) plasmid:pBEN66 DNA.
ORGANISM     artificial sequence; vectors.
REFERENCE    1 (bases 1 to 3306)
AUTHORS      Yamamoto,Y.
TITLE        Direct Submission
JOURNAL      Submitted (23-MAY-1996) to the DDBJ/EMBL/GenBank databases.
              Yoshihiro Yamamoto, Hyogo College of Medicine, Department of
              Genetics; Mukogawa-cho 1-1, Nishinomiya, Hyogo 663, Japan
              (Tel:0798-45-6587, Fax:0798-40-7639)
REFERENCE    2 (sites)
AUTHORS      Yamamoto,Y. and Furuyama,J.
TITLE        One-step disruption by circular DNA in Escherichia coli
JOURNAL      Unpublished (1996)
FEATURES     Location/Qualifiers
              1..3306
              /organism="Cloning vector pBEN66"
              /plasmid="pBEN66"
              /db_xref="taxon:47800"
              /lab_host="E.coli"
              19..36
              /note="T3 promoter"
              260..1075
              /gene="kan"
              260..1075
              /gene="kan"
              /note="derived from Tn903; kanamycin resistance gene"
              /citation=[2]
              /codon_start=1
              /transl_table=11
              /product="aminoglycoside 3'-phosphotransferase"
              /protein_id="BAAL2824.1"
              /db_xref="GI:1345434"
              /translation="MSHIQRETSCSRPLNSNMDADLYGKWARONVGSGATIVRLY
              GKPDAPLEFLKHGVSANDYTDENVRLNLTETFMPLTIKFIPTPDQAWLLTAAIP
              GKTAFOVLEETPDGSENVDAALVFLRLHSLIPVCNCFNSRVRFLAQQRMMNGL
              VDASDFDDERNGWPVEQWVKEMHKLLLPSPDSVVTHGDFSLDNLIFDEKGLGICIDVG

```

```

RVGIADRYODLAILNCLGEFSFSQKRLFKYQKIDNPDNKKLQFHLMDLDEF"
complement(2209..3069)
/genes="amp"
CDS
complement(2209..3069)
/genes="amp"
/feature="ampicillin resistance gene"
/citation=[1]
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/db_xref="GI:1345435"
/translation="MSIQHFRVALIPFFAFLPFAHPETLVKVKDAEDQLGARVGY
IEDLSNGKILESFRPEERFPMSTFKLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
YSPVTEKHTIDGTVRELCSAAITMSDNTAANLLTIIGPKELTAFLHNMGDHVTRL
DRWPELNEAIPIVNDRTMPVAMATTLKLLTGLLTSLASQQLIDMEADKVGAPL
LRSLAPGWFIADKSGAGERSGIIAALGPDGKPSRIVVITTTGSGQATMDERNRQIA
EIGASLIKHW"
BASE COUNT      854 a      800 c      790 g      862 t
ORIGIN

Query Match      47.0%; Score 70; DB 14; Length 3306;
Best Local Similarity 88.4%; Pred. No. 8.1e-12;
Matches 76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
|||||
Db 122 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 63
|||||

QY 109 GGGGGGCGCTAACTAACTAATTTTGT 134
|||||
Db 62 GGGGGGCGCGGTACCCAGCTTTTGT 37
|||||

RESULT 10
XXU35131      4144 bp      DNA      circular      SYN      26-SEP-1995
LOCUS
DEFINITION      Plasmid pBSL159 cloning vector, complete sequence.
ACCESSION      U35131
VERSION      U35131.1 GI:984907
KEYWORDS
SOURCE      Cloning vector pBSL159.
ORGANISM      Cloning vector pBSL159
              artificial sequence; vectors.
REFERENCE      1 (bases 1 to 4144)
              Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
              Improved antibiotic-resistance gene cassettes and omega elements
              for Escherichia coli vector construction and in vitro
              deletion/insertion mutagenesis
JOURNAL      Gene 160 (1), 63-67 (1995)
MEDLINE      95354958
REFERENCE      2 (bases 1 to 4144)
              Hengen, P.N.
              Direct Submission
JOURNAL      Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical
              Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
FEATURES
              source
              1..4144
              Location/Qualifiers
              /organism="Cloning vector pBSL159"
              /plasmid="pBSL159"
              /db_xref="taxon:42704"
              complement(1129..1989)
              /EC_number="3.5.2.6"
              /codon_start=1
              /transl_table=11
              /function="ampicillin resistance"
              /product="beta-lactamase"
              /protein_id="AAC3618.1"
              /db_xref="GI:984908"
              /translation="MSIQHFRVALIPFFAFLPFAHPETLVKVKDAEDQLGARVGY
              IEDLSNGKILESFRPEERFPMSTFKLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
              YSPVTEKHTIDGTVRELCSAAITMSDNTAANLLTIIGPKELTAFLHNMGDHVTRL

```

```

DRWPELNEAIPIVNDRTMPVAMATTLKLLTGLLTSLASQQLIDMEADKVGAPL
LRSLAPGWFIADKSGAGERSGIIAALGPDGKPSRIVVITTTGSGQATMDERNRQIA
EIGASLIKHW"
complement(join(3704..4144,1..219))
/codon_start=1
/transl_table=11
/feature="chloramphenicol acetyltransferase"
/protein_id="AAC3619.1"
/db_xref="GI:984909"
/translation="MEKKITGYTTVDISQWHRKEHFEAFQSVACTYVQTVOLITAF
LKTVKKNKKEYPAFIIHLARLMAHPEFRMAKDGELVINDSVHPCYTVFHEQTETF
SSLKSEYHDDRFQFLHIYSODVACYGLENLAVFKPGFIENMFVSANPWVSTFDLNV
AMDNFFAPVFTMGKYYTQGDVLMPLAIQVHVAVDGFGHVRMLNELQYCDDEWQGG
A"
BASE COUNT      1084 a      1009 c      959 g      1092 t
ORIGIN

Query Match      47.0%; Score 70; DB 14; Length 4144;
Best Local Similarity 88.4%; Pred. No. 8.3e-12;
Matches 76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
|||||
Db 2725 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2784
|||||

QY 109 GGGGGGCGCTAACTAACTAATTTTGT 134
|||||
Db 2785 GGGGGGCGCGGTACCCAGCTTTTGT 2810
|||||

RESULT 11
PRS304      4267 bp      DNA      circular      SYN      14-SEP-1995
LOCUS
DEFINITION      Yeast integrative vector pRS304 with TRP1 marker, complete
              sequence.
ACCESSION      U03436
VERSION      U03436.1 GI:416305
KEYWORDS      Cloning vector pRS304.
SOURCE      Cloning vector pRS304
              artificial sequence; vectors.
REFERENCE      1 (bases 1 to 4267)
              Sikorski, R.S. and Hieter, P.
              A system of shuttle vectors and yeast host strains designed for
              efficient manipulation of DNA in Saccharomyces cerevisiae
JOURNAL      Genetics 122 (1), 19-27 (1989)
MEDLINE      89276910
REFERENCE      2 (bases 1 to 4267)
              Stillman, D.J.
              Direct Submission
JOURNAL      Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
              and Molecular Biology, University of Utah Medical Center, Salt Lake
              City, UT 84132, USA
FEATURES
              source
              1..4267
              Location/Qualifiers
              /organism="Cloning vector pRS304"
              /db_xref="taxon:31827"
BASE COUNT      1128 a      965 c      1092 g      1082 t
ORIGIN

Query Match      47.0%; Score 70; DB 14; Length 4267;
Best Local Similarity 88.4%; Pred. No. 8.3e-12;
Matches 76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
|||||
Db 1920 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 1979
|||||

QY 109 GGGGGGCGCTAACTAACTAATTTTGT 134
|||||
Db 1980 GGGGGGCGCGGTACCCAGCTTTTGT 2005
|||||

```

LOCUS	PRS306	4373 bp	DNA	circular	SYN	14-SEP-1995					
DEFINITION	Yeast integrative vector PRS306 with URA3 marker, complete sequence.										
ACCESSION	U03438										
VERSION	U03438.1	GI:416307									
KEYWORDS											
SOURCE	Cloning vector PRS306.										
ORGANISM	Cloning vector PRS306										
REFERENCE	artificial sequence; vectors.										
AUTHORS	1 (bases 1 to 4373)										
TITLE	Sikorski,R.S. and Hieter,P.										
JOURNAL	A system of shuttle vectors and yeast host strains designed for										
MEDLINE	efficient manipulation of DNA in Saccharomyces cerevisiae										
REFERENCE	Genetics 122 (1), 19-27 (1989)										
AUTHORS	2 (bases 1 to 4373)										
TITLE	Stillman,D.J.										
JOURNAL	Direct Submission										
FEATURES	Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral										
SOURCE	and Molecular Biology, University of Utah Medical Center, Salt Lake										
BASE COUNT	1181 a	987 c	1106 g	1099 t							
ORIGIN	Location/Qualifiers										
	1..4373										
	/organism="Cloning vector PRS306"										
	/db_xref="taxon:31829"										
Query Match	47.0%; Score 70; DB 14; Length 4373;										
Best Local Similarity	88.4%; Pred. No. 8.3e-12;										
Matches	76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;										
QY	49	AGTGGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACTCGAG	108								
Db	2026	AGTGGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACTCGAG	2085								
QY	109	GGGGGGCGCTAACTAACTAATTTGTT	134								
Db	2086	GGGGGGCGGCTACCCAGCTTTGTT	2111								
RESULT 14											
PRS303											
LOCUS	PRS303	4443 bp	DNA	circular	SYN	14-SEP-1995					
DEFINITION	Yeast integrative vector PRS303 with HIS3 marker, complete sequence.										
ACCESSION	U03435										
VERSION	U03435.1	GI:416304									
KEYWORDS											
SOURCE	Cloning vector PRS303.										
ORGANISM	Cloning vector PRS303										
REFERENCE	artificial sequence; vectors.										
AUTHORS	1 (bases 1 to 4443)										
TITLE	Sikorski,R.S. and Hieter,P.										
JOURNAL	A system of shuttle vectors and yeast host strains designed for										
MEDLINE	efficient manipulation of DNA in Saccharomyces cerevisiae										
REFERENCE	Genetics 122 (1), 19-27 (1989)										
AUTHORS	2 (bases 1 to 4443)										
TITLE	Stillman,D.J.										
JOURNAL	Direct Submission										
FEATURES	Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral										
SOURCE	and Molecular Biology, University of Utah Medical Center, Salt Lake										
BASE COUNT	1149 a	1048 c	1109 g	1137 t							
ORIGIN	Location/Qualifiers										
	1..4443										
	/organism="Cloning vector PRS303"										

```

RESULT 12
XXU35136      4289 bp      DNA      circular      SYN      26-SEP-1995
LOCUS
DEFINITION   Plasmid pBSL97 cloning vector, complete sequence.
ACCESSION    U35136
VERSION      U35136.1      GI:984923
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS      Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
TITLE        Improved antibiotic-resistance gene cassettes and omega elements
              for Escherichia coli vector construction and in vitro
              deletion/insertion mutagenesis
JOURNAL
MEDLINE      Gene 160 (1), 53-67 (1995)
REFERENCE    95354958
AUTHORS      2 (bases 1 to 4289)
TITLE        Direct Submission
JOURNAL      Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical
              Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
FEATURES
source
    /organism="Cloning vector pBSL97"
    /plasmid="pBSL97"
    /db_xref="taxon:42709"
    complement(804..1598)
    /codon_start=1
    /transl_table=1
    /product="neomycin phosphotransferase"
    /protein_id="AAC53629.1"
    /db_xref="GI:984924"
    /translation="MIEQQLHAGSPAANVERLFGVDNAQQIGTCSDAVFRLSAQRG
    PVLFVKLTUSGALNEIQDEARLSWLTGVPCAVLDVVTAGRWMLLIGVEPGQDL
    LSSHLAPAEKVSIMADAMRRLHTLDPATCFDQAKHRIERARTRMKGLVPQDDLDE
    EHQLGAPAEFLARLARMPPGDELVVTGHDACLNTIMVNGRFSGFIGDGRGLGVADRY
    ODIALTRDIAEELGGEWADRFLVLYGIAAPDSQRIAFYRLDLEFF"
CDS
    /cds_number="3.5..2.6"
    /transl_table=1
    /function="ampicillin resistance"
    /product="beta-lactamase"
    /protein_id="AAC53630.1"
    /db_xref="GI:984925"
    /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVYVKDAEDQLGARVGY
    IELDSNGRIELFEFRPEELSPKKMTSTKVLCCGAVLSRIDAGQQLGRRIHYSONDLVE
    YSPVTEKHTLDGTVTAEELSAATMSDNTAANLLTIIGTKELTAFLNMGGDHVTRL
    DWREPENAFIPNDERDTPMVMATTLRKLTGELLTLASRQQLIDMWEADKVGKPL
    LRSALPAGFIADKSGAGERSGIIAALGPDGKPSRIIVYITGSGQATMDENRQII
    EIGASLIKHW"
BASE COUNT   1024 a 1157 c 1111 g 997 t
ORIGIN

Query Match      47.0%; Score 70; DB 14; Length 4289;
Best Local Similarity 88.4%; Pred. No. 8.3e-12;
Matches 76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY  49  AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
      |||||
Db   2014 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2073
      |||||

QY  109 GGGGGGCGCTAACTAACTATTTGTT 134
      |||||
Db   2074 GGGGGGCGCGGTACCCAGCTTTGTT 2099
      |||||

```

Query Match 47.0%; Score 70; DB 14; Length 4443;
Best Local Similarity 88.4%; Pred. No. 8.3e-12;
Matches 76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Search completed: May 29, 2000, 21:35:38
Job time: 36870 sec

QY 49 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 108
DB 2106 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 2165
QY 109 GGGGGGCTAACTAATTTTGT 134
DB 2166 GGGGGGCTAACTAATTTTGT 2191

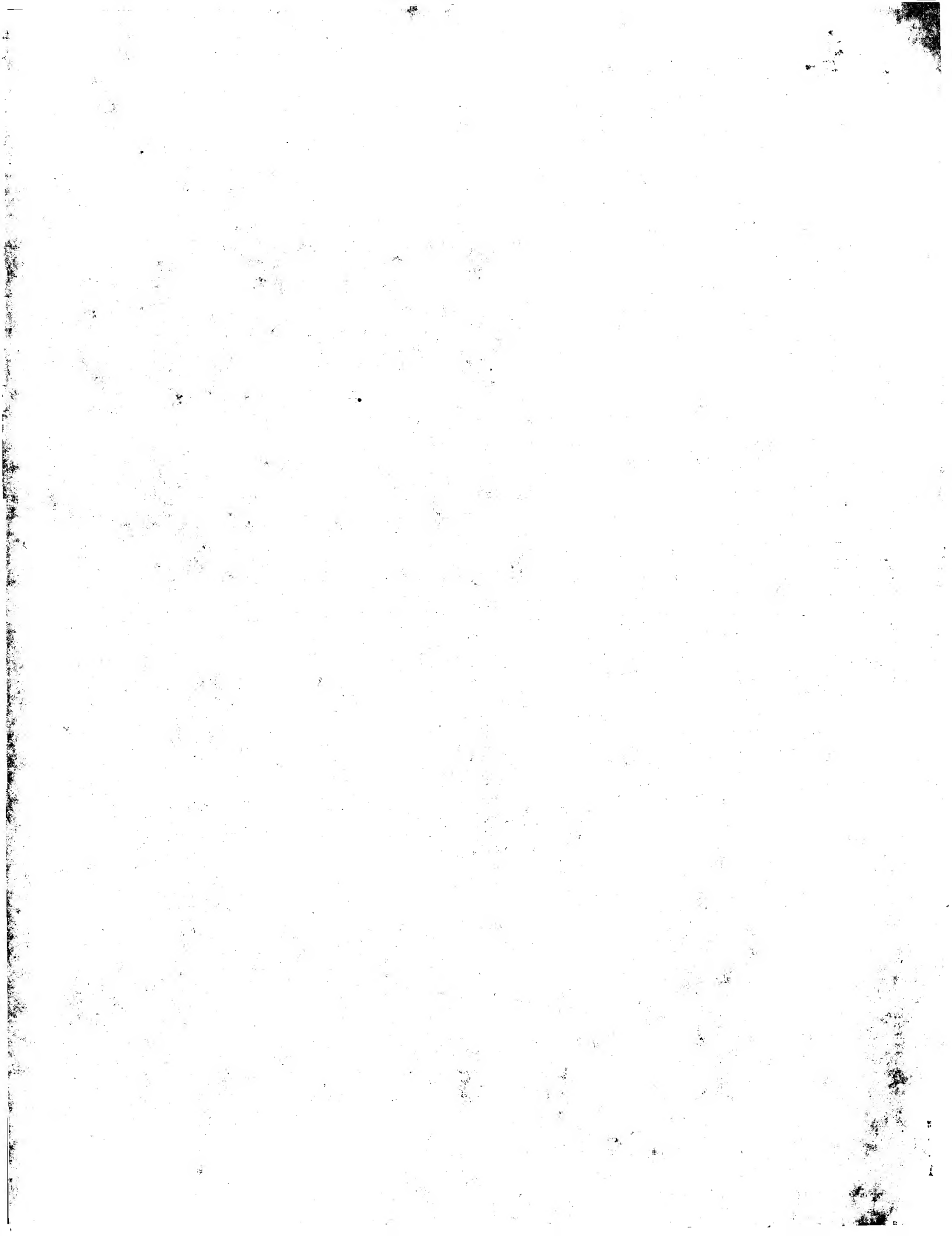
RESULT 15

ASAJ5326/c ASAJ5326 4670 bp DNA circular SYN 08-FEB-1999
LOCUS PGAI(+) KS positive selection cloning vector glts gene.
DEFINITION
ACCESSION AJ005326
VERSION AJ005326.1 GI:4028984
KEYWORDS glts gene; glutamate permease.
SOURCE synthetic construct.
ORGANISM
1 (bases 1 to 4670)
REFERENCE
AUTHORS Gal, J.
TITLE Direct Submission
JOURNAL Submitted (03-APR-1998) Gal, J., Institute for Biotechnology, Bay
Zoltan Foundation for Applied Research, Szeged, Derkovits fasor 2.,
6726, HUNGARY
REFERENCE
AUTHORS Gal, J., Szekeres, S., Schnell, R., Pongor, S., Simoncsits, A. and
Kalman, M.
TITLE A positive selection cloning system based on the glts gene of
Escherichia coli
JOURNAL Anal. Biochem. 266 (2), 235-238 (1999)
MEDLINE 99107575
FEATURES
Location/Qualifiers
1..4670
/organism="synthetic construct"
/db_xref="taxon:32630"
852..2231
/gene="glts"
852..2231
/gene="glts"
/codon_start=1
/transl_table=11
/product="glutamate permease"
/protein_id="CAA06473.1"
/db_xref="GI:4028985"
/translation="MFHLDLTATLVAAITLLGRKLVHSFLKKYTIPEPVAGLL
VALALLVKKMGWEVNFDMSLRDPMLAFAITGNNANIASLRAGRWGIFLIVV
GLIVMNAIGIDMLGLAGSITLGGHGTGAWSKLFIERYGTNATEV
AMACATFGLVGLIGPVARQLTGTGKNSWPGPPSRSTVSLISNSCPDPLV
LERPPRWSNSPYSEYVYLVKHTTPNGIPDDQEVPTAFKPDVGRMITSLVLI
ETIALTICLTGVKIVAOILLACTAFELPFCVLFVGLNSGLSIMGFYRVERAVS
VLGNVSLSLFLMALMGLKELASLALPMLAVQIENALYALFVETWRMGKNYD
AAVLAAGHCGFGGATPTAIAANQATTEFGPSHNAFLVVPVNGAFFIDIYNALVIK
YLMPIFAG"

BASE COUNT 1040 a 1165 c 1231 g 1234 t
ORIGIN

Query Match 47.0%; Score 70; DB 14; Length 4670;
Best Local Similarity 88.4%; Pred. No. 8.4e-12;
Matches 76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 108
DB 1501 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 1442
QY 109 GGGGGGCTAACTAATTTTGT 134
DB 1441 GGGGGGCTAACTAATTTTGT 1416



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
C 1	71.2	47.8	259	1	Q87664	Mouse azoospermia	
C 2	70	47.0	501	1	T04866	Nucleotide analog	
C 3	68	45.6	651	1	X20513	Polynucleotide seq	
C 4	68	45.6	752	1	V31294	E. coli J96 pathog	
C 5	68	45.6	1091	1	T75006	Human endogenous r	
C 6	68	45.6	1122	1	T75005	Human endogenous r	
C 7	68	45.6	3792	1	Q48463	Plasmid pg+host4 c	
C 8	68	45.6	4226	1	T39485	Human steroidogene	
C 9	68	45.6	5234	1	Q48464	Plasmid pg+host5 c	
C 10	68	45.6	6722	1	Q48465	Plasmid pg+host6 c	
C 11	68	45.6	10529	1	V09028	Maize fluoery2 gene	
C 12	68	45.6	12814	1	X24730	Swedish-FAD APP ta	
C 13	68	45.6	15692	1	X24731	London-FAD APP tar	
C 14	68	45.6	15692	1	X24732	Swedish/London-FAD	
C 15	68	45.6	15701	1	X24733	Swedish-FAD APP713	
C 16	67.8	45.5	545	1	V68808	Human endogenous r	
C 17	67.4	45.2	2840	1	Q28664	BDAT. cDNA encodin	
C 18	66.2	44.4	3481	1	X02815	DE19731274 Seq ID	
C 19	66.2	44.4	3810	1	X02813	DE19731274 Seq ID	
C 20	65.4	43.9	5356	1	T43794	Plasmid pRIPAT (r	
C 21	65	43.6	84	1	X02800	E. coli biotin DNA	
C 22	65	43.6	3465	1	X02814	DE19731274 Seq ID	
C 23	65	43.6	3794	1	X02812	DE19731274 Seq ID	
C 24	65	43.6	6596	1	V57377	Maize female-prefe	
C 25	64.6	43.4	545	1	T75010	Human endogenous r	
C 26	63.4	42.6	685	1	V10190	Stealth virus nucl	
C 27	63.4	42.6	685	1	V12003	Stealth virus plas	
C 28	60	40.3	4164	1	T04575	Plasmid pAT-1 sequ	
C 29	60	40.3	4164	1	V22271	PAR-1 (pS0544). DN	
C 30	60	40.3	4164	1	V69740	Nucleotide sequenc	
C 31	60	40.3	5178	1	T49876	pret-splice. Nucle	
C 32	58.8	39.5	2973	1	V64254	Plasmid pKS varian	
C 33	58.8	39.5	3198	1	T93702	Candida Carbol gen	

DT 28-JAN-1996 (first entry)
DE Nucleotide analogue treated with calf intestinal alkaline phosphatase.
KW Nucleotide analogue; alkaline phosphatase; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..501
FT /*tag= a
FT /note= "std. IUPAC codes used"
PN WO9524505-A.
PD 14-SEP-1995.
PF 07-MAR-1995; G00490.
PR 08-MAR-1994; EP-301636.
PA (AMSH) AMERSHAM INT PLC.
PI Reeve MA, Robinson PS;
DR WPI; 95-328290/42.
PT Modification of residual fluorescence labelled nucleotide analogues - to
PT prevent migration in electrophoretic sequencing gel and interference with
PT base calling of DNA chains
PS Example; Fig 1; 18pp; English.
CC The invention concerns the modification of residual fluorescence
CC labelled nt analogues to prevent migration in electrophoretic
CC sequencing gel and interference with base calling of DNA chains. The
CC modification involves the use of a phosphatase enzyme to remove at
CC least one 5'-phosphate gp. The nts were prepared using ABI Amplitaq
CC dye-terminating kit. The template used was 1 microg. of M13 mp8. The
CC primer was M13 universal primer. Samples were recovered. Some were
CC subjected to calf intestinal alkaline phosphatase digestion. This
CC includes R04866. R04867 was a control nt. which was not treated with
CC alkaline phosphatase. The example shows that alkaline phosphatase
CC treatment causes removal of the dye-terminator artefacts and allows
CC for accurate base calling with the ABI analysis software.
CC Sequence 501 BP; 98 A; 136 C; 140 G; 121 T;
SQ

Query Match 47.0%; Score 70; DB 1; Length 501;
Best Local Similarity 88.4%; Pred. No. 2.4e-15;
Matches 76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 108
DB 56 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 115

QY 109 GGGGGGCCCTAACTAACTAATTTTGT 134
DB 116 GGGGGGCCCGGTACCCAGCTTTTGT 141

RESULT 3
ID X20513 standard; DNA; 651 BP.
AC X20513;
AT 05-MAY-1999 (first entry)
DE Polynucleotide sequence from the genome of Treponema pallidum.
KW treponema pallidum infection; syphilis; Borrelia infection; animal;
KW enzyme production; ds.
OS treponema pallidum.
PN WO9859034-A2.
PD 30-DEC-1998.
PF 23-JUN-1998; U13041.
PR 24-JUN-1997; US-050667.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Fraser CM;
DR WPI; 99-081273/07.
PT New isolated treponema pallidum nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention
PT and therapy of T. pallidum infections, particularly syphilis
PS Claim 1; Page 257; 1150pp; English.
CC X20500-21243 represent polynucleotide sequences from the genome of
CC treponema pallidum. The sequences can be used for detection,
CC diagnosis, characterisation, prevention and therapy for T. pallidum
CC infections, particularly syphilis. They can also be used for detecting
CC diseases related to Borrelia infections in animals, and for the
CC production of biosynthetic products such as enzymes.

SQ Sequence 651 BP; 180 A; 167 C; 152 G; 149 T;

Query Match 45.6%; Score 68; DB 1; Length 651;
Best Local Similarity 100.0%; Pred. No. 1.3e-14;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 108
DB 77 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 136

QY 109 GGGGGGCC 116
DB 137 GGGGGGCC 144

RESULT 4
V31294/c
ID V31294 standard; DNA; 752 BP.
AC V31294;
AT 01-OCT-1998 (first entry)
DE E. coli J96 pathogenicity island contig #108.
KW PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pheR;
KW PAI V; pheV; vaccine; protective immune response; ds.
OS Escherichia coli.
PN WO9822575-A2.
PD 28-MAY-1998.
PF 21-NOV-1997; U21347.
PR 14-OCT-1997; US-061953.
PR 22-NOV-1996; US-031626.
PA (HUMA-) HUMAN GENOME SCI INC.
PA (UYWI-) UNIV WISCONSIN.
PI Choi GH, Dillon PJ, Welch RA;
PI WPI; 98-312461/27.
DR New isolated uropathogenic E. coli nucleotide sequences - used to
DR develop products for the detection of pathogenic E. coli and to
DR elicit an immune response to pathogenic E. coli
PS Claim 21; Page 221-222; 250pp; English.
CC This sequence represents a E. coli strain J96 contig containing
CC pathogenicity island (PAI) sequences, and represents a nucleic acid
CC molecule of the invention. PAIs are large fragments of DNA which comprise
CC pathogenicity determinants. The sequences of the invention are taken from
CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)
CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at
CC approximately 94 min (at pheR) on the E. coli chromosome and is
CC approximately 160 kb in size. Antibodies specific to the proteins encoded
CC by the PAI open reading frames of the invention can be used in kits to
CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit
CC a protective immune response in an animal to the uropathogenic E. coli
CC strain J96.
SQ Sequence 752 BP; 162 A; 213 C; 203 G; 172 T;

Query Match 45.6%; Score 68; DB 1; Length 752;
Best Local Similarity 100.0%; Pred. No. 1.4e-14;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 108
DB 73 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 14

QY 109 GGGGGGCC 116
DB 13 GGGGGGCC 6

RESULT 5
ID T75006/c
AC T75006;
AT 06-OCT-1997 (first entry)
DE Human endogenous retroviral sequence 6.
DE Breast cancer; tumour; B18a1; prognosis; diagnosis; vaccine; ss.
KW

```
OS Human retrovirus.
PN WO9725431-A1.
PD 17-JUL-1997.
PF 10-JAN-1997: U00398.
PR 10-JAN-1996: US-587329.
PA (CORI-) CORIXA CORP.
PI Frudakis TN, Smith JM;
WPI: 97-384982/35.
PT Endogenous human tumour-associated retroviral element, B18Ag1 - used
for the prognosis, diagnosis and monitoring of human cancers,
especially breast cancer.
PS Claim 10: Page 31-32: 74pp; English.
CC Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14
and 11-22 (T75003-10) were obtd. by screening human genomic
libraries using human breast tumour-associated retroviral element
B18Ag1 (see also T75002) as probe. These non-contiguous sequences
lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the
retrovirus genome (see also T75001). B18Ag1 and the other
retroviral sequences can be used in genetic vaccines and for the
prognosis, diagnosis and monitoring of human breast cancer.
SQ Sequence 1091 BP; 79 A; 350 C; 97 G; 248 T;

Query Match 45.6%; Score 68; DB 1: Length 1091;
Best Local Similarity 100.0%; Pred. No. 1.5e-14; Indels 0; Gaps 0;
Matches 68; Conservative 0; Mismatches 0;

QY 49 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
DB 138 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 79
QY 109 GGGGGGCC 116
DB 78 GGGGGGCC 71

RESULT 6
ID T75005/c
AC T75005;
DT 06-OCT-1997 (first entry)
DE Human endogenous retroviral sequence 3.
KW Breast cancer; tumour; B18Ag1; prognosis; diagnosis; vaccine; ss.
OS Human retrovirus.
PN WO9725431-A1.
PD 17-JUL-1997.
PF 10-JAN-1997: U00398.
PR 10-JAN-1996: US-587329.
PA (CORI-) CORIXA CORP.
PI Frudakis TN, Smith JM;
WPI: 97-384982/35.
PT Endogenous human tumour-associated retroviral element, B18Ag1 - used
for the prognosis, diagnosis and monitoring of human cancers,
especially breast cancer.
PS Claim 10: Page 30-31: 74pp; English.
CC Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14
and 11-22 (T75003-10) were obtd. by screening human genomic
libraries using human breast tumour-associated retroviral element
B18Ag1 (see also T75002) as probe. These non-contiguous sequences
lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the
retrovirus genome (see also T75001). B18Ag1 and the other
retroviral sequences can be used in genetic vaccines and for the
prognosis, diagnosis and monitoring of human breast cancer.
SQ Sequence 1122 BP; 260 A; 316 C; 185 G; 279 T;

Query Match 45.6%; Score 68; DB 1: Length 1122;
Best Local Similarity 100.0%; Pred. No. 1.6e-14; Indels 0; Gaps 0;
Matches 68; Conservative 0; Mismatches 0;

QY 49 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
DB 109 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 50
```

```
QY 109 GGGGGGCC 116
DB 49 GGGGGGCC 42

RESULT 7
QY 49 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
DB 3542 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 3483
QY 109 GGGGGGCC 116
DB 3482 GGGGGGCC 3475

RESULT 8
ID T39485/c
AC T39485;
DT 21-MAY-1997 (first entry)
DE Human steroidogenesis acute regulatory protein genomic DNA.
KW Human; steroidogenesis; acute regulatory protein; hSTAR; analysis;
mutation; detection; prenatal; genetic defect; congenital; protein;
lipoid adrenal hyperplasia; treatment; prevention; gene;
KW Replacement therapy; hypercholesterolaemia; ds.
OS Homo sapiens.
PN WO9629338-A1.
PD 26-SEP-1996.
PF 22-MAR-1996: U03896.
PR 23-MAR-1995: US-410540.
PA (REGC) UNIV CALIFORNIA.
PA (UYPE-) UNIV PENNSYLVANIA.
PI Lin D, Miller WL, Strauss JF;
WPI: 96-443130/44.
PT Isolated human steroidogenesis acute regulatory protein gene - used
```

for detection of mutation(s) of this gene that cause congenital

PT lipid adrenal hyperplasia

PS Claim 1; Pages 23-25; 89pp; English.

CC The present sequence encodes the human steroidogenesis acute

CC regulatory protein (hSTAR). The hSTAR gene can be analysed for

CC mutations to detect (e.g. prenatally) genetic defects associated

CC with congenital lipid adrenal hyperplasia (CAH), or its

CC transmission to children. CAH can be treated by protein or gene

CC replacement therapy, which can also be used to prevent or treat

CC hypercholesterolaemia.

CC A human adrenal cortex cDNA library was screened with a mouse STAR

CC probe to isolate a 1.6 kb insert, including an ORF for a 285

CC residue protein. When it was cloned into pSPORT and expressed in

CC COS-1 cells cotransfected with pP450scab and pADX, it increased the

CC level of pregnenolone synthesis from cholesterol or

CC 20-alpha-hydroxycholesterol.

CC 20-alpha-hydroxycholesterol.

SQ Sequence 4226 BP; 940 A; 1132 C; 1144 G; 984 T;

Query Match 45.6%; Score 68; DB 1; Length 4226;

Best Local Similarity 100.0%; Pred. No. 2.4e-14; Indels 0; Gaps 0;

Matches 68; Conservative 0; Mismatches 0;

QY 49 AGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGACCTCGAG 108

DB 4148 AGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGACCTCGAG 4207

QY 109 GGGGGGCC 116

DB 4208 GGGGGGCC 4215

RESULT 9

Q48464/c

ID Q48464 standard; DNA; 5234 BP.

AC Q48464;

DT 18-MAR-1994 (first entry)

DE Plasmid pG+host5 containing Ts replication system.

KW Temperature sensitive replication; antibiotic resistance marker gene;

KW site-specific recombination; chromosomal integration; inactivation;

KW heterologous gene expression; thermosensitive plasmid; ds.

OS Synthetic.

FH Key

FT misc_feature 36..1496

FT Location/Qualifiers

FT /tag= a

FT /standard_name= ORI

FT /note= "Origin of replication from pBR322"

FT 2640..4383

FT /tag= b

FT /note= "from pGK12 (derived from pWV01)"

FT 4384..4786

FT /tag= c

FT /note= "from PUB110"

FT 4787..5234

FT /tag= d

FT /note= "from pSK"

FT W09318164-A.

PN 16-SEP-1993.

PD 12-MAR-1992; F00248.

PF (INRG) INRA INST NAT RECH AGRONOMIQUE.

PA Gruss A, Maguin E;

PI WPI; 93-303478/38.

PT New bacterial plasmid contg. heat sensitive replication system -

PT and marker gene, opt. capable of chromosomal integration, used to

PT inactivate specific gene or introduce heterologous gene

PS Disclosure; Fig 11; 73pp; French.

CC Plasmid pGK12 (Appl. Environ. Microbiol., 48; 726 (1984)) contg. two

CC antibiotic resistance marker genes was subjected to mutagenesis with

CC hydroxylamine. A heat-stable mutant was isolated (coding for a heat-

CC sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and

CC the 3340bp fragment lacking the Cm resistance gene was ligated to a

CC 445bp PvuII fragment of pBluescript SK+ containing a multicloning

CC site, T7 and T3 promoters and sequencing primer binding sites. The

CC resulting plasmid was designated pVE6004 (or pG+host4 - Q48463).

CC To facilitate cloning in E.coli, the Avai-EcoRI fragment of pBR322

CC (containing the origin of replication and the ampicillin resistance

CC gene) was inserted into NsiI-cleaved, blunt-ended pG+host4 to give

CC pG+host6 (Q48466). The pBR322 ORI allows maintenance of the plasmid in

CC E.coli at 37 deg.C. while the heat-sensitive bacteria.

CC E.coli at 28 deg.C in gram-positive bacteria.

SQ Sequence 6722 BP; 1961 A; 1278 C; 1503 G; 1980 T;

Query Match 45.6%; Score 68; DB 1; Length 6722;

Best Local Similarity 100.0%; Pred. No. 2.8e-14; Indels 0; Gaps 0;

Matches 68; Conservative 0; Mismatches 0;

QY 49 AGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGACCTCGAG 108

DB 6472 AGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGACCTCGAG 6413

QY 109 GGGGGGCC 116

DB 6412 GGGGGGCC 6405

CC site, T7 and T3 promoters and sequencing primer binding sites. The

CC resulting plasmid was designated pVE6004 (or pG+host4 - Q48463).

CC To facilitate cloning in E.coli, the 1.4kb Avai-AluNI fragment of

CC pBR322 (containing the origin of replication) was inserted into

CC NsiI-cleaved pG+host4 to give pG+host5 (Q48464). The pBR322 ORI

CC allows maintenance of the plasmid in E.coli at 37 deg.C. while the

CC heat-sensitive ORI allows maintenance at 28 deg.C in gram-positive

CC bacteria.

SQ Sequence 5234 BP; 1569 A; 973 C; 1133 G; 1559 T;

Query Match 45.6%; Score 68; DB 1; Length 5234;

Best Local Similarity 100.0%; Pred. No. 2.6e-14; Indels 0; Gaps 0;

Matches 68; Conservative 0; Mismatches 0;

QY 49 AGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGACCTCGAG 108

DB 4984 AGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGACCTCGAG 4925

QY 109 GGGGGGCC 116

DB 4924 GGGGGGCC 4917

RESULT 10

Q48465/c

ID Q48465 standard; DNA; 6722 BP.

AC Q48465;

DT 18-MAR-1994 (first entry)

DE Plasmid pG+host6 containing Ts replication system.

KW Temperature sensitive replication; antibiotic resistance marker gene;

KW site-specific recombination; chromosomal integration; inactivation;

KW heterologous gene expression; thermosensitive plasmid; ds.

OS Synthetic.

PN W09318164-A.

PD 16-SEP-1993.

PF 12-MAR-1992; F00248.

PR (INRG) INRA INST NAT RECH AGRONOMIQUE.

PI Gruss A, Maguin E;

PI WPI; 93-303478/38.

PT New bacterial plasmid contg. heat sensitive replication system -

PT and marker gene, opt. capable of chromosomal integration, used to

PT inactivate specific gene or introduce heterologous gene

PS Disclosure; Fig 11; 73pp; French.

CC Plasmid pGK12 (Appl. Environ. Microbiol., 48; 726 (1984)) contg. two

CC antibiotic resistance marker genes was subjected to mutagenesis with

CC hydroxylamine. A heat-stable mutant was isolated (coding for a heat-

CC sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and

CC the 3340bp fragment lacking the Cm resistance gene was ligated to a

CC 445bp PvuII fragment of pBluescript SK+ containing a multicloning

CC site, T7 and T3 promoters and sequencing primer binding sites. The

CC resulting plasmid was designated pVE6004 (or pG+host4 - Q48463).

CC To facilitate cloning in E.coli, the Avai-EcoRI fragment of pBR322

CC (containing the origin of replication and the ampicillin resistance

CC gene) was inserted into NsiI-cleaved, blunt-ended pG+host4 to give

CC pG+host6 (Q48466). The pBR322 ORI allows maintenance of the plasmid in

CC E.coli at 37 deg.C. while the heat-sensitive bacteria.

CC E.coli at 28 deg.C in gram-positive bacteria.

SQ Sequence 6722 BP; 1961 A; 1278 C; 1503 G; 1980 T;

Query Match 45.6%; Score 68; DB 1; Length 6722;

Best Local Similarity 100.0%; Pred. No. 2.8e-14; Indels 0; Gaps 0;

Matches 68; Conservative 0; Mismatches 0;

KW homologous recombination; transgenic mouse; transgenic animal;
 KW animal model; Alzheimer's disease; vector; pMTI-2453;
 KW London-FAD APP; mutation; ss.
 OS Homo sapiens.
 OS Synthetic.

PH Key Location/Qualifiers
 FT mat_peptide 4807..5151
 FT /*tag= a
 FT /standard_name= "Swedish-FAD"
 FT /note= "encodes W97998"
 FT mutation replace(4990,"")
 FT /*tag= b
 FT /standard_name= "London-FAD"
 FT 8223..9023
 FT mat_peptide /*tag= c
 FT /note= "encodes W97998"

PN WO9909150-A1.
 PD 25-FEB-1999.
 PF 18-AUG-1997; U14507.
 PF 18-AUG-1997; WO-U14507.
 PR (FARB) BAYER CORP.
 PA Wirak DO;
 PI WPI; 99-181029/15.
 DR P-PSDB; W97998, W97998.
 DR Modification of target nucleic acids - by homologous recombination,
 PT used particularly for introducing a humanised amyloid precursor
 PT protein gene into rodents for producing models of Alzheimer's
 PT disease
 PT Example; Page 104-113; 209pp; English.
 CC This is the nucleotide sequence of London-FAD APP targeting
 CC vector pMTI-2453. The invention provides a novel gene targeting
 CC strategy that facilitates the introduction of one or more specific
 CC mutations into any gene in a single double reciprocal homologous
 CC recombination step. The method has been used particularly for
 CC introducing a humanised amyloid precursor protein (APP) gene into
 CC rodents for producing animal models of Alzheimer's disease (AD).
 CC 4 Independent lines of transgenic mice (lines ES5007, ES5103,
 CC ES5401 and ES5403) have been created using the gene targeting
 CC technique applied to embryonic stem cells. In each line, the mouse
 CC APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
 CC amino acid residues 666-770 of APP770 were encoded by human CDNA
 CC sequences instead of mouse genomic exons (exons 16-18). Within
 CC these residues, only 3 amino acid differences exist between the
 CC mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
 CC Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
 CC containing a humanised beta-amyloid domain. In line ES5401,
 CC created using vector pMTI-2453, the London mutation, i.e.
 CC V(717)I, was also introduced.
 CC Sequence 15692 BP; 2423 A; 2730 C; 2786 G; 2541 T;

Query Match 45.6%; Score 68; DB 1; Length 15692;
 Best Local Similarity 100.0%; Pred. No. 3.7e-14;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
 DB 12765 AGTGGATCCCCGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 12824
 QY 109 GGGGGGCC 116
 DB 12825 GGGGGGCC 12832

RESULT 14
 X24732
 ID X24732 standard; DNA; 15692 BP.
 AC X24732;
 DT 21-JUN-1999 (first entry)
 DE Swedish/London-FAD APP targeting vector pMTI-2454.
 KW Amyloid precursor protein; APP; human; gene targeting;
 KW homologous recombination; transgenic mouse; transgenic animal;
 KW animal model; Alzheimer's disease; vector; pMTI-2454;

KW London-FAD APP; Swedish-FAD APP; mutation; ss.
 OS Homo sapiens.
 OS Synthetic.
 PH Key Location/Qualifiers
 FT mat_peptide 4807..5151
 FT /*tag= a
 FT /note= "encodes W98000"
 FT mutation replace(4849,"")
 FT /*tag= b
 FT /standard_name= "Swedish-FAD"
 FT mutation replace(4989,"")
 FT /standard_name= "London-FAD"
 FT 8223..9023
 FT mat_peptide /*tag= b
 FT /note= "encodes W97998"

PN WO9909150-A1.
 PD 25-FEB-1999.
 PF 18-AUG-1997; U14507.
 PF 18-AUG-1997; WO-U14507.
 PR (FARB) BAYER CORP.
 PA Wirak DO;
 PI WPI; 99-181029/15.
 DR P-PSDB; W97998, W98000.
 DR Modification of target nucleic acids - by homologous recombination,
 PT used particularly for introducing a humanised amyloid precursor
 PT protein gene into rodents for producing models of Alzheimer's
 PT disease
 PT Example; Page 114-123; 209pp; English.
 CC This is the nucleotide sequence of Swedish/London-FAD APP targeting
 CC vector pMTI-2454. The invention provides a novel gene targeting
 CC strategy that facilitates the introduction of one or more specific
 CC mutations into any gene in a single double reciprocal homologous
 CC recombination step. The method has been used particularly for
 CC introducing a humanised amyloid precursor protein (APP) gene into
 CC rodents for producing animal models of Alzheimer's disease (AD).
 CC 4 Independent lines of transgenic mice (lines ES5007, ES5103,
 CC ES5401 and ES5403) have been created using the gene targeting
 CC technique applied to embryonic stem cells. In each line, the mouse
 CC APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
 CC amino acid residues 666-770 of APP770 were encoded by human CDNA
 CC sequences instead of mouse genomic exons (exons 16-18). Within
 CC these residues, only 3 amino acid differences exist between the
 CC mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
 CC Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
 CC containing a humanised beta-amyloid domain. In line ES5103,
 CC created using vector pMTI-2454, the London mutation, i.e.
 CC V(717)I, and the Swedish mutation, i.e. KM(670,671)NL, were
 CC introduced. The targetted Swedish/London-FAD m/hAPP gene expressed
 CC m/hAPP protein at levels approaching those observed for mouse APP
 CC in brain.
 CC Sequence 15692 BP; 2423 A; 2731 C; 2783 G; 2542 T;

Query Match 45.6%; Score 68; DB 1; Length 15692;
 Best Local Similarity 100.0%; Pred. No. 3.7e-14;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
 DB 12765 AGTGGATCCCCGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 12824
 QY 109 GGGGGGCC 116
 DB 12825 GGGGGGCC 12832

RESULT 15
 X24733
 ID X24733 standard; DNA; 15701 BP.
 AC X24733;
 DT 21-JUN-1999 (first entry)
 DE Swedish-FAD APP713 targeting vector pMTI-2454.
 KW Amyloid precursor protein; APP; human; gene targeting;

KW homologous recombination; transgenic mouse; transgenic animal;
KW animal model; Alzheimer's disease; vector; pMTI-2455;
KW Swedish-FAD APP713; mutation; ss.

OS Homo sapiens.

OS Synthetic.

FH Key Location/Qualifiers
FT mat_peptide 4807..4983

FT FT /*tag= a

FT FT /note= "encodes W98001"

FT FT replace(4835,"")

FT FT /*tag= b

FT FT /standard_name= "Swedish-FAD"

FT FT replace(4981,"")

FT FT /*tag= c

FT FT /standard_name= APP713stop

FT FT 8232..9032

FT FT /*tag= d

FT FT /note= "encodes W97998"

FN W09909150-A1.

PD 25-FEB-1999.

PF 18-AUG-1997; U14507.

PR 18-AUG-1997; WO-U14507.

PA (FARB) BAYER CORP.

PI Wirak DO;

DR WPI: 99-181029/15.

DR P-PSDB; W97998, W98001.

PT Modification of target nucleic acids - by homologous recombination,
PT used particularly for introducing a humanised amyloid precursor
PT protein gene into rodents for producing models of Alzheimer's
PT disease

PS Example; Page 124-133; 209pp; English.

CC This is the nucleotide sequence of Swedish-FAD APP713 targeting
CC vector pMTI-2455. The invention provides a novel gene targeting
CC strategy that facilitates the introduction of one or more specific
CC mutations into any gene in a single double reciprocal homologous
CC recombination step. The method has been used particularly for
CC introducing a humanised amyloid precursor protein (APP) gene into
CC rodents for producing animal models of Alzheimer's disease (AD).

CC 4 Independent lines of transgenic mice (lines ES5007, ES5103,
CC ES5401 and ES5403) have been created using the gene targeting
CC technique applied to embryonic stem cells. In each line, the mouse
CC APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC amino acid residues 666-770 of APP770 were encoded by human cDNA
CC sequences instead of mouse genomic exons (exons 16-18). Within
CC these residues, only 3 amino acid differences exist between the
CC mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
CC containing a humanised beta-amyloid domain. In line ES5215,
CC created using vector pMTI-2455, the Swedish mutation. i.e.
CC KM(670,681)NL, was also introduced, and also a stop codon (T to
CC stop at position 714).

CC Sequence 15701 BP; 2470 A; 2675 C; 2845 G; 2497 T;

SQ

Query Match

45.6%; Score 68; DB 1; Length 15701;

Best Local Similarity 100.0%; Pred. No. 3.7e-14;

Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 49 AGTGGATCCCCCGGTCGAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108

Db 12774 AGTGGATCCCCCGGTCGAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 12833

OY 109 GGGGGGCC 116

Db 12834 GGGGGGCC 12841

Search completed: May 29, 2000, 21:58:33
Job time: 38181 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 22:09:05 ; Search time 621.83 Seconds
(without alignments)
31.146 Million cell updates/sec

Title: US-08-935-377-8

Perfect score: 149
Sequence: 1 GCCCAAAATTGAAACAACTA.....TGTGTTTGTGGGCGCCGCC 149

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgnl_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgnl_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgnl_6/ptodata/1/ina/5C_COMB.seq.*
4: /cgnl_6/ptodata/1/ina/5D_COMB.seq.*
5: /cgnl_6/ptodata/1/ina/6_COMB.seq.*
6: /cgnl_6/ptodata/1/ina/PCTUS_COMB.seq.*
7: /cgnl_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	68	45.6	3792	4	US-08-992-334-1
C 2	68	45.6	3792	5	US-08-302-752-1
C 3	68	45.6	4016	2	US-08-410-540-3
C 4	68	45.6	5234	4	US-08-992-334-2
C 5	68	45.6	5234	5	US-08-302-752-2
C 6	68	45.6	6722	4	US-08-992-334-3
C 7	68	45.6	6722	5	US-08-302-752-3
C 8	67.8	45.5	793	3	US-08-367-101-121
C 9	67.8	45.5	793	4	US-08-592-541-121
C 10	65.4	43.9	9318	3	US-08-793-610-6
C 11	63.4	42.6	685	2	US-08-463-115-56
C 12	63.4	42.6	685	2	US-08-465-388-56
C 13	60	40.3	4164	1	US-08-204-675-1
C 14	60	40.3	4164	3	US-08-660-754-1
C 15	60	40.3	4164	4	US-08-796-364-1
C 16	60	40.3	4164	6	PCT-US95-02520-1
C 17	60	40.3	5178	3	US-08-474-169-2
C 18	56	37.6	88	1	US-08-144-602B-15
C 19	56	37.6	1023	1	US-08-198-446B-1
C 20	56	37.6	1023	3	US-08-870-693-1
C 21	56	37.6	2150	1	US-08-198-446B-10
C 22	56	37.6	2150	3	US-08-870-693-10
C 23	55	36.9	834	3	US-08-967-101-113
C 24	55	36.9	834	4	US-08-592-541-113
C 25	51.6	34.6	3341	3	US-08-868-577-18
C 26	51.2	34.4	1200	1	US-08-011-598B-3

ALIGNMENTS

RESULT 1

US-08-992-334-1/c

; Sequence 1, Application US/08992334

; Patent No. 5919678

; GENERAL INFORMATION:

; APPLICANT: Gruss, Alexandra

; APPLICANT: Mayuin, Emmanuelle

; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE

; TITLE OF INVENTION: PLASMID

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Christie Parker & Hale, LLP

; STREET: 350 West Colorado Boulevard, Suite 500

; CITY: Pasadena

; STATE: California

; COUNTRY: United States

; ZIP: 91105

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/992,334

; FILING DATE: 17-DEC-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/302,752

; FILING DATE: 24-DEC-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/FR93/00248

; FILING DATE: 12-MAR-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: FR 92/03034

; FILING DATE: 13-MAR-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Prout, D. Bruce

; REGISTRATION NUMBER: 20958

; REFERENCE/DOCKET NUMBER: C93:31779

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (626) 795-9900

; TELEFAX: (626) 577-8800

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3792 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: both

; TOPOLOGY: circular

; MOLECULE TYPE: DNA (genomic)

Sequence 3, Appl1
Sequence 3, Appl1
Sequence 116, App
Sequence 116, App
Sequence 5, Appl1
Sequence 15, Appl1
Sequence 8, Appl1
Sequence 8, Appl1
Sequence 62, Appl1
Sequence 1, Appl1
Sequence 1, Appl1
Sequence 3, Appl1
Sequence 35, Appl1
Sequence 35, Appl1
Sequence 3, Appl1
Sequence 1, Appl1
Sequence 4, Appl1
Sequence 5, Appl1

```
;
;
; HYPOTHEICAL: NO
; ANTI-SENSE: YES
; IMMEDIATE SOURCE:
; CLONE: pg+host4
;
US-08-992-334-1

Query Match          45.68; Score 68; DB 4; Length 3792;
Best Local Similarity 100.0%; Pred. No. 6.6e-15;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGATCCCGGGCTCGAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
      |||||||
DB 3542 AGTGATCCCGGGCTCGAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 3483
      |||||||

QY 109 GGGGGGCC 116
      |||||||
DB 3482 GGGGGGCC 3475

RESULT 2
US-08-302-752-1/c
; Sequence 1, Application US/08302752
; Patent No. 6025190
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: THERMOSENSIBLE PLASMID
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,752
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9203034
; FILING DATE: 13-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR/93/00248
; FILING DATE: 12-MAR-1993
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3792 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-08-302-752-1

Query Match          45.68; Score 68; DB 5; Length 3792;
Best Local Similarity 100.0%; Pred. No. 6.6e-15;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGATCCCGGGCTCGAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
      |||||||
DB 3542 AGTGATCCCGGGCTCGAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 3483
      |||||||

QY 109 GGGGGGCC 116
      |||||||
DB 3482 GGGGGGCC 3475

RESULT 3
US-08-410-540-3
; Sequence 3, Application US/08410540
; Patent No. 5807678
; GENERAL INFORMATION:
; APPLICANT: Miller, Walter L.
; APPLICANT: Lin, Dong
```

```
;
;
; APPLICANT: Strauss III, Jerome F.
; TITLE OF INVENTION: IDENTIFICATION OF GENE MUTATIONS
; TITLE OF INVENTION: ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94306-2155
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/410,540
; FILING DATE: 23-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Neeley, Richard L.
; REGISTRATION NUMBER: 30,092
; REFERENCE/DOCKET NUMBER: UCAL-238/000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415 853 5070
; TELEFAX: 415 857 0663
; TELEX: 380816COOLEYPA
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1098..1283
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1620..1733
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2047..2174
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2267..2425
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2567..2751
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2828..2921
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3031..3765
; FEATURE:
; NAME/KEY:
; LOCATION: 1433..1434
; OTHER INFORMATION: /note= "interruption of sequence data"
; FEATURE:
; NAME/KEY:
; LOCATION: 2208..2209
; OTHER INFORMATION: /note= "interruption of sequence data"
; FEATURE:
; NAME/KEY:
; LOCATION: 2781..2782
; OTHER INFORMATION: /note= "interruption of sequence data"
; FEATURE:
; NAME/KEY:
; LOCATION: 2956..2957
```

OTHER INFORMATION: /note= "interruption of sequence data"
US-08-410-540-3

Query Match 45.6%; Score 68; DB 2; Length 4016;
Best Local Similarity 100.0%; Pred. No. 6.7e-15;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 108
|||||
Db 3938 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 3997
|||||
QY 109 GGGGGGCC 116
|||||
Db 3998 GGGGGGCC 4005

RESULT 4
US-08-992-334-2/c
; Sequence 2, Application US/08992334
; Patent No. 5919678
; GENERAL INFORMATION:
; APPLICANT: Gruss, Alexandra
; APPLICANT: Maguin, Emmanuelle
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE
; TITLE OF INVENTION: PLASMID
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christie Parker & Hale, LLP
; STREET: 350 West Colorado Boulevard, Suite 500
; CITY: Pasadena
; STATE: California
; COUNTRY: United States
; ZIP: 91105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/992.334
FILING DATE: 17-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302.752
FILING DATE: 24-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00248
FILING DATE: 12-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR FR92/03034
FILING DATE: 13-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Prout, D. Bruce
REGISTRATION NUMBER: 20958
REFERENCE/DOCKET NUMBER: C93:31779
TELECOMMUNICATION INFORMATION:
TELEPHONE: (626) 795-9900
TELEFAX: (626) 577-8800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5234 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-992-334-2

Query Match 45.6%; Score 68; DB 4; Length 5234;
Best Local Similarity 100.0%; Pred. No. 7.4e-15;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 108
|||||
Db 4984 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 4925
|||||
QY 109 GGGGGGCC 116
|||||
Db 4924 GGGGGGCC 4917

RESULT 5
US-08-302-752-2/c
; Sequence 2, Application US/08302752
; Patent No. 6025190
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: THERMOSENSIBLE PLASMID
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302.752
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9203034
; FILING DATE: 13-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR/93/00248
; FILING DATE: 12-MAR-1993
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5234 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-302-752-2

Query Match 45.6%; Score 68; DB 5; Length 5234;
Best Local Similarity 100.0%; Pred. No. 7.4e-15;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 108
|||||
Db 4984 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 4925
|||||
QY 109 GGGGGGCC 116
|||||
Db 4924 GGGGGGCC 4517

RESULT 6
US-08-992-334-3/c
; Sequence 3, Application US/08992334
; Patent No. 5919678
; GENERAL INFORMATION:
; APPLICANT: Gruss, Alexandra
; APPLICANT: Maguin, Emmanuelle
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE
; TITLE OF INVENTION: PLASMID
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christie Parker & Hale, LLP
; STREET: 350 West Colorado Boulevard, Suite 500
; CITY: Pasadena
; STATE: California
; COUNTRY: United States
; ZIP: 91105
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/992,334
FILING DATE: 17-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302,752
FILING DATE: 24-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00248
FILING DATE: 12-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/03034
FILING DATE: 13-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Prout, D. Bruce
REGISTRATION NUMBER: 20958
REFERENCE/DOCKET NUMBER: C93:31779
TELECOMMUNICATION INFORMATION:
TELEPHONE: (626) 795-9900
TELEFAX: (626) 577-8800
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6722 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-992-334-3

Query Match 45.6%; Score 68; DB 4; Length 6722;
Best Local Similarity 100.0%; Pred. No. 8.1e-15;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
Db 6472 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 6413

QY 109 GGGGGGCC 116
Db 6412 GGGGGGCC 6405

RESULT 7
US-08-302-752-3/c
Sequence 3, Application US/08302752
Patent No. 6025190
GENERAL INFORMATION:
APPLICANT: THERMOSENSIBLE PLASMID
TITLE OF INVENTION: THERMOSENSIBLE PLASMID
NUMBER OF SEQUENCES: 3
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/302,752
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9203034
FILING DATE: 13-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO FR/93/00248
FILING DATE: 12-MAR-1993
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6722 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-302-752-3

Query Match 45.6%; Score 68; DB 5; Length 6722;
Best Local Similarity 100.0%; Pred. No. 8.1e-15;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
Db 6472 AGTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 6413

QY 109 GGGGGGCC 116
Db 6412 GGGGGGCC 6405

RESULT 8
US-08-967-101-121/c
Sequence 121, Application US/08967101
Patent No. 5840540
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,101
FILING DATE: 10-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,541

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 793 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-967-101-121

Query Match 45.5%; Score 67.8; DB 3; Length 793;
Best Local Similarity 86.2%; Pred. No. 4.5e-15;
Matches 75; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 48 GAGTGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 107
Db 120 GGGCGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 61

QY 108 GGGGGGGCCTACTACTAATTTTGT 134
Db 60 GGGGGGGCCCGGTACCGCTTTTGT 34

RESULT 9

US-08-592-541-121/c
; Sequence 121, Application US/08592541
; Patent No. 5986054
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROWMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-592-541-121

Query Match 45.5%; Score 67.8; DB 4; Length 793;
Best Local Similarity 86.2%; Pred. No. 4.5e-15;
Matches 75; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 48 GAGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 107
Db 120 GGGCGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGA 61
QY 108 GGGGGGGCCTACTACTAATTTTGT 134
Db 60 GGGGGGGCCCGGTACCGCTTTTGT 34

RESULT 10

US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BADM, Christopher
; APPLICANT: STOCKING-HARRIS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaigo, Marmelstein, Murray & Oram LLP

; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4 4 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-08-793-610-6

Query Match 43.9%; Score 65.4; DB 3; Length 9318;
Best Local Similarity 98.5%; Pred. No. 7.6e-14;
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 50 GTGATFCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 109
Db 5654 GGGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 5713
QY 110 GGGGGGCC 116
Db 5714 GGGGGGCC 5720

RESULT 11

US-08-463-115-56
; Sequence 56, Application US/08463115
; Patent No. 5703221
; GENERAL INFORMATION:
; APPLICANT: WILLIAM JOHN MARTIN
; TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
; TITLE OF INVENTION: AND RELATED VACCINES
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible

```
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,115
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/157,811
; FILING DATE: No. 5703221ember 23, 1993
; APPLICATION NUMBER: 07/887,502
; FILING DATE: May 22, 1992
; APPLICATION NUMBER: 07/704,814
; FILING DATE: May 23, 1991
; APPLICATION NUMBER: 07/763,039
; FILING DATE: September 20, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 213/301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 685 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION:
;
US-08-463-115-56

Query Match 42.6%; Score 63.4; DB 2; Length 685;
Best Local Similarity 94.1%; Pred. No. 1.5e-13;
Matches 64; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
Db 149 AGTGGATCCCCGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 208

QY 109 GGGGGGCC 116
Db 209 GGGGGGNC 216

RESULT 12
US-08-465-388-56
; Sequence 56, Application US/08465388
; Patent No. 5753488
; GENERAL INFORMATION:
; APPLICANT: WILLIAM JOHN MARTIN
; TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
; TITLE OF INVENTION: AND RELATED VACCINES
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; COUNTRY: California
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/465,388
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/157,811
; FILING DATE: No. 5753488ember 23, 1993
; APPLICATION NUMBER: 07/887,502
; FILING DATE: May 22, 1992
; APPLICATION NUMBER: 07/704,814
; FILING DATE: May 23, 1991
; APPLICATION NUMBER: 07/763,039
; FILING DATE: September 20, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 213/300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 685 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION:
;
US-08-465-388-56

Query Match 42.6%; Score 63.4; DB 2; Length 685;
Best Local Similarity 94.1%; Pred. No. 1.5e-13;
Matches 64; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
Db 149 AGTGGATCCCCGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 208

QY 109 GGGGGGCC 116
Db 209 GGGGGGNC 216

RESULT 13
US-08-204-675-1
; Sequence 1, Application US/08204675
; Patent No. 5671170
; GENERAL INFORMATION:
; APPLICANT: Devine, Scott E.
; APPLICANT: Boeke, Jef D.
; APPLICANT: Braiterman, Lelita T.
; TITLE OF INVENTION: In Vitro Transposition of Artificial
; TITLE OF INVENTION: Transposons
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie, and Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,675
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 435
```



```
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.45501
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202.508.9100
; TELEFAX: 202.508.9299
; TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4164 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: PAT-1
;
US-08-204-675-1

Query Match 40.3%; Score 60; DB 1; Length 4164;
Best Local Similarity 100.0%; Pred. No. 4.6e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
DB 2043 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102

RESULT 14
US-08-660-754-1
; Sequence 1, Application US/08660754
; Patent No. 5843772
; GENERAL INFORMATION:
; APPLICANT: Devine, Scott E.
; APPLICANT: Boeke, Jef D.
; APPLICANT: Braiterman, Lelita T.
; TITLE OF INVENTION: In Vitro Transposition of Artificial
; TITLE OF INVENTION: Transposons
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie, and Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/660,754
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,675
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.45501
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202.508.9100
; TELEFAX: 202.508.9299
; TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4164 base pairs
; TYPE: nucleic acid
```

```
;
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: PAT-1
;
US-08-660-754-1

Query Match 40.3%; Score 60; DB 3; Length 4164;
Best Local Similarity 100.0%; Pred. No. 4.6e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
DB 2043 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102

RESULT 15
US-08-796-364-1
; Sequence 1, Application US/08796364
; Patent No. 5968785
; GENERAL INFORMATION:
; APPLICANT: Devine, Scott E.
; APPLICANT: Boeke, Jef D.
; APPLICANT: Braiterman, Lelita T.
; TITLE OF INVENTION: In Vitro Transposition of Artificial
; TITLE OF INVENTION: Transposons
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie, and Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/796,364
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,675
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.45501
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202.508.9100
; TELEFAX: 202.508.9299
; TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4164 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: PAT-1
;
US-08-796-364-1

Query Match 40.3%; Score 60; DB 4; Length 4164;
Best Local Similarity 100.0%; Pred. No. 4.6e-12;
```

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 AGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
|||||
Db 2043 AGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102
|||||

Search completed: May 29, 2000, 22:09:10
Job time: 38751 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2000, 09:48:31 ; Search time 2276.24 Seconds
(without alignments)
292.062 Million cell updates/sec

Title: US-08-935-377-8
Perfect score: 149
Sequence: 1 GGCCAAAATTGAAACAACTA.....TTGTTTTTGGCCCGGCC 149

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5142629 seqs, 2230885800 residues
Total number of hits satisfying chosen parameters: 10285240

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA.*
1: /cgnl_6/ptodata/1/pna/US092B_COMB.seq.*
2: /cgnl_6/ptodata/1/pna/US092C_COMB.seq.*
3: /cgnl_6/ptodata/1/pna/US092D_COMB.seq.*
4: /cgnl_6/ptodata/1/pna/US092E_COMB.seq.*
5: /cgnl_6/ptodata/1/pna/US093A_COMB.seq.*
6: /cgnl_6/ptodata/1/pna/US093B_COMB.seq.*
7: /cgnl_6/ptodata/1/pna/US093C_COMB.seq.*
8: /cgnl_6/ptodata/1/pna/US093D_COMB.seq.*
9: /cgnl_6/ptodata/1/pna/US093E_COMB.seq.*
10: /cgnl_6/ptodata/1/pna/US093F_COMB.seq.*
11: /cgnl_6/ptodata/1/pna/US094A_COMB.seq.*
12: /cgnl_6/ptodata/1/pna/US094B_COMB.seq.*
13: /cgnl_6/ptodata/1/pna/US094C_COMB.seq.*
14: /cgnl_6/ptodata/1/pna/US094D_COMB.seq.*
15: /cgnl_6/ptodata/1/pna/US094E_COMB.seq.*
16: /cgnl_6/ptodata/1/pna/US094F_COMB.seq.*
17: /cgnl_6/ptodata/1/pna/US095A_COMB.seq.*
18: /cgnl_6/ptodata/1/pna/US095B_COMB.seq.*
19: /cgnl_6/ptodata/1/pna/US095C_COMB.seq.*
20: /cgnl_6/ptodata/1/pna/US095D_COMB.seq.*
21: /cgnl_6/ptodata/1/pna/US095E_COMB.seq.*
22: /cgnl_6/ptodata/1/pna/US095F_COMB.seq.*
23: /cgnl_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgnl_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgnl_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgnl_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgnl_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgnl_6/ptodata/1/pna/US096F_COMB.seq.*
29: /cgnl_6/ptodata/1/pna/US096G_COMB.seq.*
30: /cgnl_6/ptodata/1/pna/US096H_COMB.seq.*
31: /cgnl_6/ptodata/1/pna/US096I_COMB.seq.*
32: /cgnl_6/ptodata/1/pna/US096J_COMB.seq.*
33: /cgnl_6/ptodata/1/pna/US096K_COMB.seq.*
34: /cgnl_6/ptodata/1/pna/US096L_COMB.seq.*
35: /cgnl_6/ptodata/1/pna/US096M_COMB.seq.*
36: /cgnl_6/ptodata/1/pna/US096N_COMB.seq.*
37: /cgnl_6/ptodata/1/pna/US096O_COMB.seq.*
38: /cgnl_6/ptodata/1/pna/US096P_COMB.seq.*
39: /cgnl_6/ptodata/1/pna/US096Q_COMB.seq.*
40: /cgnl_6/ptodata/1/pna/US096R_COMB.seq.*
41: /cgnl_6/ptodata/1/pna/US096S_COMB.seq.*
42: /cgnl_6/ptodata/1/pna/US096T_COMB.seq.*
43: /cgnl_6/ptodata/1/pna/US096U_COMB.seq.*

44: /cgnl_6/ptodata/1/pna/US092B_COMB.seq.*
45: /cgnl_6/ptodata/1/pna/US092C_COMB.seq.*
46: /cgnl_6/ptodata/1/pna/US092D_COMB.seq.*
47: /cgnl_6/ptodata/1/pna/US092E_COMB.seq.*
48: /cgnl_6/ptodata/1/pna/US093A_COMB.seq.*
49: /cgnl_6/ptodata/1/pna/US093B_COMB.seq.*
50: /cgnl_6/ptodata/1/pna/US093C_COMB.seq.*
51: /cgnl_6/ptodata/1/pna/US093D_COMB.seq.*
52: /cgnl_6/ptodata/1/pna/US093E_COMB.seq.*
53: /cgnl_6/ptodata/1/pna/US093F_COMB.seq.*
54: /cgnl_6/ptodata/1/pna/US094A_COMB.seq.*
55: /cgnl_6/ptodata/1/pna/US094B_COMB.seq.*
56: /cgnl_6/ptodata/1/pna/US094C_COMB.seq.*
57: /cgnl_6/ptodata/1/pna/US094D_COMB.seq.*
58: /cgnl_6/ptodata/1/pna/US094E_COMB.seq.*
59: /cgnl_6/ptodata/1/pna/US094F_COMB.seq.*
60: /cgnl_6/ptodata/1/pna/US095A_COMB.seq.*
61: /cgnl_6/ptodata/1/pna/US095B_COMB.seq.*
62: /cgnl_6/ptodata/1/pna/US095C_COMB.seq.*
63: /cgnl_6/ptodata/1/pna/US095D_COMB.seq.*
64: /cgnl_6/ptodata/1/pna/US095E_COMB.seq.*
65: /cgnl_6/ptodata/1/pna/US095F_COMB.seq.*
66: /cgnl_6/ptodata/1/pna/US096A_COMB.seq.*
67: /cgnl_6/ptodata/1/pna/US096B_COMB.seq.*
68: /cgnl_6/ptodata/1/pna/US096C_COMB.seq.*
69: /cgnl_6/ptodata/1/pna/US096D_COMB.seq.*
70: /cgnl_6/ptodata/1/pna/US096E_COMB.seq.*
71: /cgnl_6/ptodata/1/pna/US096F_COMB.seq.*
72: /cgnl_6/ptodata/1/pna/US096G_COMB.seq.*
73: /cgnl_6/ptodata/1/pna/US096H_COMB.seq.*
74: /cgnl_6/ptodata/1/pna/US096I_COMB.seq.*
75: /cgnl_6/ptodata/1/pna/US096J_COMB.seq.*
76: /cgnl_6/ptodata/1/pna/US096K_COMB.seq.*
77: /cgnl_6/ptodata/1/pna/US096L_COMB.seq.*
78: /cgnl_6/ptodata/1/pna/US096M_COMB.seq.*
79: /cgnl_6/ptodata/1/pna/US096N_COMB.seq.*
80: /cgnl_6/ptodata/1/pna/US096O_COMB.seq.*
81: /cgnl_6/ptodata/1/pna/US096P_COMB.seq.*
82: /cgnl_6/ptodata/1/pna/US096Q_COMB.seq.*
83: /cgnl_6/ptodata/1/pna/US096R_COMB.seq.*
84: /cgnl_6/ptodata/1/pna/US096S_COMB.seq.*
85: /cgnl_6/ptodata/1/pna/US096T_COMB.seq.*
86: /cgnl_6/ptodata/1/pna/US096U_COMB.seq.*
87: /cgnl_6/ptodata/1/pna/US096V_COMB.seq.*
88: /cgnl_6/ptodata/1/pna/US096W_COMB.seq.*
89: /cgnl_6/ptodata/1/pna/US096X_COMB.seq.*
90: /cgnl_6/ptodata/1/pna/US096Y_COMB.seq.*
91: /cgnl_6/ptodata/1/pna/US096Z_COMB.seq.*
92: /cgnl_6/ptodata/1/pna/US097A_COMB.seq.*
93: /cgnl_6/ptodata/1/pna/US097B_COMB.seq.*
94: /cgnl_6/ptodata/1/pna/US097C_COMB.seq.*
95: /cgnl_6/ptodata/1/pna/US097D_COMB.seq.*
96: /cgnl_6/ptodata/1/pna/US097E_COMB.seq.*
97: /cgnl_6/ptodata/1/pna/US097F_COMB.seq.*
98: /cgnl_6/ptodata/1/pna/US097G_COMB.seq.*
99: /cgnl_6/ptodata/1/pna/US097H_COMB.seq.*
100: /cgnl_6/ptodata/1/pna/US097I_COMB.seq.*
101: /cgnl_6/ptodata/1/pna/US097J_COMB.seq.*
102: /cgnl_6/ptodata/1/pna/US097K_COMB.seq.*
103: /cgnl_6/ptodata/1/pna/US097L_COMB.seq.*
104: /cgnl_6/ptodata/1/pna/US097M_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	149	100.0	149	28	US-08-935-377-8
2	138	92.6	150	28	US-08-935-377-9

```

137 91.9 148 28 US-08-935-377-7
131 87.9 145 28 US-08-935-377-6
70 47.0 224 13 US-08-377-383-4
70 47.0 224 18 US-08-693-573-4
69.4 46.6 838 40 US-09-123-912-91
69.4 46.6 858 43 US-09-221-107-91
69.4 46.6 858 46 US-09-285-479-91
68 45.6 147 38 US-09-072-433-31
68 45.6 157 63 US-60-029-960-595
68 45.6 245 54 US-09-411-999-19130
68 45.6 381 49 US-09-332-782-8570
68 45.6 381 57 US-09-515-694-8570
68 45.6 393 54 US-09-411-999-13627
68 45.6 405 49 US-09-332-782-8523
68 45.6 405 57 US-09-515-694-8523
68 45.6 420 49 US-09-332-782-8595
68 45.6 420 57 US-09-515-694-8595
68 45.6 422 49 US-09-332-782-9309
68 45.6 422 57 US-09-515-694-9309
68 45.6 423 54 US-09-411-999-34199
68 45.6 433 42 US-09-179-473-2334
68 45.6 433 49 US-09-328-351-2334
68 45.6 442 42 US-09-179-473-2495
68 45.6 442 49 US-09-328-351-2495
68 45.6 448 42 US-09-328-351-2255
68 45.6 448 49 US-09-328-351-2255
68 45.6 454 42 US-09-179-473-2100
68 45.6 454 49 US-09-328-351-2100
68 45.6 454 49 US-09-328-351-2360
68 45.6 454 32 US-08-978-259-108
68 45.6 752 64 US-60-031-626-108
68 45.6 752 82 US-60-126-265-936
68 45.6 2958 37 US-09-040-961-6
68 45.6 2958 41 US-09-130-115-6
68 45.6 2961 15 US-08-446-935-6
68 45.6 543 17 US-08-583-392-10
68 45.6 543 17 US-08-583-392A-10
68 45.6 545 17 US-08-587-329-10
68 45.6 545 19 US-08-700-014-10
68 45.6 545 23 US-08-838-762-10
68 45.6 545 34 US-08-991-788-10
68 45.6 545 103 US-09-429-755-10

ALIGNMENTS

Sequence 7, Appl1
Sequence 6, Appl1
Sequence 4, Appl1
Sequence 91, Appl1
Sequence 91, Appl1
Sequence 91, Appl1
Sequence 31, Appl1
Sequence 595, Appl1
Sequence 19130, A
Sequence 8370, Ap
Sequence 8370, Ap
Sequence 13627, A
Sequence 8523, Ap
Sequence 8523, Ap
Sequence 8595, Ap
Sequence 8595, Ap
Sequence 9309, Ap
Sequence 9309, Ap
Sequence 34199, A
Sequence 2334, Ap
Sequence 2495, Ap
Sequence 2495, Ap
Sequence 2255, Ap
Sequence 2255, Ap
Sequence 2100, Ap
Sequence 2100, Ap
Sequence 2360, Ap
Sequence 2360, Ap
Sequence 108, Ap
Sequence 108, Ap
Sequence 996, Ap
Sequence 6, Appl1
Sequence 6, Appl1
Sequence 10, Appl1
Sequence 10, Appl1
Sequence 10, Appl1
Sequence 10, Appl1
Sequence 10, Appl1
Sequence 10, Appl1

US-08-935-377-8
NAME: Steffe, Eric K
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 149 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-935-377-8

Query Match 100.0%; Score 149; DB 28; Length 149;
Best Local Similarity 100.0%; Pred. No. 5.9e-40;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACAAAATTGAAAACTAGATCTATTTATTCACGCGCGCCGATGAGTGATCCCCC 60
|||||
Db 1 GGCACAAAATTGAAAACTAGATCTATTTATTCACGCGCGCCGATGAGTGATCCCCC 60
|||||

QY 61 GGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCCTAAC 120
|||||
Db 61 GGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCCTAAC 120
|||||

QY 121 TAACTAATTTTGTGTTTGTGGCGCGGCC 149
|||||
Db 121 TAACTAATTTTGTGTTTGTGGCGCGGCC 149
|||||

RESULT 2
US-08-935-377-9
; Sequence 9, Application US/08935377
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; TITLE OF INVENTION: T Cells Specific for Target Antigens and
; TITLE OF INVENTION: Vaccines Based Thereon
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,377
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-935-377-9
```



```
STATE: DC
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 720Kb storage
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1 or ASCII editors
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/377,383
FILING DATE: 25-JAN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP6971/1994
FILING DATE: 26-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Glenn J. Perry
REGISTRATION NUMBER: 28458
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)861-3000
TELEFAX: (202)822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid Synthetic DNA
US-08-377-383-4

Query Match 47.0%; Score 70; DB 13; Length 224;
Best Local Similarity 88.4%; Pred. No. 2e-13;
Matches 76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGGCTCGAGAAATTCGATATCAAGCTTATCGATACCGTGCAGCTCGAG 108
DB 87 AGTGGATCCCCGGGCTCGAGAAATTCGATATCAAGCTTATCGATACCGTGCAGCTCGAG 146
QY 109 GGGGGCCCTAACTAACTAATTTGTT 134
DB 147 GGGGGCCCGGTACCCAGCTTTGTT 172

RESULT 6
US-08-693-573-4
; Sequence 4, Application US/08693573
; GENERAL INFORMATION:
; APPLICANT: HOSOI, Shigeru,
; APPLICANT: FUKAMI, Tadashi,
; APPLICANT: KOJIMA, Makiko
; TITLE OF INVENTION: Method of Determining Base Sequence of
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY CUSHMAN
; ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: Ninth Floor, 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 720Kb storage
; COMPUTER: IBM PC/XT/AT compatibles
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1 or ASCII editors
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,573
; FILING DATE: 07-AUG-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/377,383
```

Nucleic

```
FILING DATE: 25-JAN-1995
APPLICATION NUMBER: JP6971/1994
FILING DATE: 26-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Thomas G. Wiseman
REGISTRATION NUMBER: 35046
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)861-3000
TELEFAX: (202)822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid Synthetic DNA
US-08-693-573-4

Query Match 47.0%; Score 70; DB 18; Length 224;
Best Local Similarity 88.4%; Pred. No. 2e-13;
Matches 76; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGGCTCGAGAAATTCGATATCAAGCTTATCGATACCGTGCAGCTCGAG 108
DB 87 AGTGGATCCCCGGGCTCGAGAAATTCGATATCAAGCTTATCGATACCGTGCAGCTCGAG 146
QY 109 GGGGGCCCTAACTAACTAATTTGTT 134
DB 147 GGGGGCCCGGTACCCAGCTTTGTT 172

RESULT 7
US-09-123-912-91
; Sequence 91, Application US/09123912A
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongcong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C1
; CURRENT APPLICATION NUMBER: US/09/123,912A
; CURRENT FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: 09/040,802
; PRIOR FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (570)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (591)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (655)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (664)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (667)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (683)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (711)
; OTHER INFORMATION: Where n is a, c, g or t
```

```
; NAME/KEY: modified_base
; LOCATION: (759)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (760)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (765)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (777)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (787)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (792)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (794)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (801)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (804)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (809)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (817)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (820)
; OTHER INFORMATION: Where n is a, c, g or t
;
; US-09-123-912-91
;
; Query Match 46.6% Score 69.4; DB 40; Length 858;
; Best Local Similarity 76.6% Pred. No. 4.5e-13;
; Matches 85; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
;
; Qy 6 AAAATTGAAACTAGATCTATTATTCACGCGCGGCCCATGATGATCCCGGGCT 65
;      ||||| | |||| | |||| | ||||| |||||
; Db 313 ataactcaccaggagattcaacgaattccaccacactggactagtgtatcccccgggct 372
;
; Qy 66 GCAGGAATTCGATCACTGATCTATTCGATACCGTCGACCTCGAGGGGGGCC 116
;      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; Db 373 gcaggaattcgatcaagcttatcgatccgacgcgagggggggcc 423
;
; RESULT 8
; US-09-221-107-91
; Sequence 91, Application US/09221107
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C2
; CURRENT APPLICATION NUMBER: US/09/221.107
; CURRENT FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (570)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; US-09-221-107-91
```

```
Query Match          46.6%; Score 69.4; DB 43; Length 858;
Best Local Similarity 76.6%; Pred. No. 4.5e-13;
Matches 85; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 6 AAAAAAAGTATGATCTATTATTGACGGCGCCGATGAGTGGATCCCGGGCT 65
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 313 ataattcaccaggagattcaagaattccaccactgagtagtgatcccccggt 372

QY 66 GCAGGAATTCGATATCAAGTTATCGATACCGTCGACCTCGAGGGGGGCC 116
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 373 gcaggaattcgatcaagcttatcgataccgtcgacctgagggggggcc 423

RESULT 9
US-09-285-479-91
; Sequence 91, Application US/09285479
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: LUNG CANCER
; FILE REFERENCE: 210121.455C3
; CURRENT APPLICATION NUMBER: US/09/285,479
; CURRENT FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(858)
; OTHER INFORMATION: n = A,T,C or G
US-09-285-479-91
```

```
Query Match          46.6%; Score 69.4; DB 46; Length 858;
Best Local Similarity 76.6%; Pred. No. 4.5e-13;
Matches 85; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 6 AAAAAAAGTATGATCTATTATTGACGGCGCCGATGAGTGGATCCCGGGCT 65
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 313 ataattcaccaggagattcaagaattccaccactgagtagtgatcccccggt 372

QY 66 GCAGGAATTCGATATCAAGTTATCGATACCGTCGACCTCGAGGGGGGCC 116
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 373 gcaggaattcgatcaagcttatcgataccgtcgacctgagggggggcc 423

RESULT 10
US-09-072-433-31/c
; Sequence 31, Application US/09072433
; GENERAL INFORMATION:
; APPLICANT: Coschigano, Peter W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: BIOREMEDIATION
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Cartoll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,433
; FILING DATE: 04-MAY-1998
; CLASSIFICATION: 514
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Cartoll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OHU-03344
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 147 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..147
US-09-072-433-31

Query Match          45.6%; Score 68; DB 38; Length 147;
Best Local Similarity 100.0%; Pred. No. 8.2e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 ACTGGATCCCCGGGCTGCAGGAATTCGATATCAAGTTATCGATACCGTCGACCTCGAG 108
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 107 ACTGGATCCCCGGGCTGCAGGAATTCGATATCAAGTTATCGATACCGTCGACCTCGAG 48

QY 109 GGGGGGCC 116
   ||||| ||||| ||||| |||||
Db 47 GGGGGGCC 40

RESULT 11
US-60-029-960-595/c
; Sequence 595, Application US/60029960
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 1649
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/029,960
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 595:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
```


US-60-029-960-595

Query Match 45.6%; Score 68; DB 63; Length 157;
Best Local Similarity 100.0%; Pred. No. 8.3e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGATCCCCGGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
|||||

Db 93 AGTGATCCCCGGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 34
|||||

QY 109 GGGGGGCC 116
|||||

Db 33 GGGGGGCC 26
|||||

RESULT 12

US-09-411-999-19130

; Sequence 19130, Application US/09411999

; GENERAL INFORMATION:

; APPLICANT: Havukkala, Ilkka

; APPLICANT: Shenk, Michael Andrew

; TITLE OF INVENTION: Polynucleotides isolated from plants and

; FILE REFERENCE: 1023U

; CURRENT APPLICATION NUMBER: US/09/411,999

; CURRENT FILING DATE: 1999-10-04

; NUMBER OF SEQ ID NOS: 37897

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 19130

; LENGTH: 245

; TYPE: DNA

; ORGANISM: Pinus radiata

US-09-411-999-19130

Query Match 45.6%; Score 68; DB 54; Length 245;

Best Local Similarity 100.0%; Pred. No. 9.4e-13;

Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGATCCCCGGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
|||||

Db 30 agtggatccccgggctcgaggaaatcgatatcaagcttatcgatcacgtcgacctcgag 89
|||||

QY 109 GGGGGGCC 116
|||||

Db 90 ggggggcc 97
|||||

RESULT 13

US-09-332-782-8570

; Sequence 8570, Application US/09332782

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE REFERENCE: 20411-754CON1

; CURRENT APPLICATION NUMBER: US/09/332,782

; CURRENT FILING DATE: 1999-06-14

; EARLIER APPLICATION NUMBER: US 09/181,430

; EARLIER FILING DATE: 1998-10-28

; NUMBER OF SEQ ID NOS: 21027

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 8570

; LENGTH: 381

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(381)

; OTHER INFORMATION: n = A,T,C or G

US-09-332-782-8570

Query Match 45.6%; Score 68; DB 49; Length 381;

Best Local Similarity 100.0%; Pred. No. 1.1e-12;

Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGATCCCCGGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
|||||

Db 97 agtggatccccgggctcgaggaaatcgatatcaagcttatcgatcacgtcgacctcgag 156
|||||

QY 109 GGGGGGCC 116
|||||

Db 157 ggggggcc 164
|||||

RESULT 14

US-09-515-694-8570

; Sequence 8570, Application US/09515694

; GENERAL INFORMATION:

; APPLICANT: Arterburn, Matthew

; APPLICANT: Asghari, Vida

; APPLICANT: Bluh, Linda

; APPLICANT: Cheung, Patrick

; APPLICANT: Damavandi, Simin

; APPLICANT: Dickson, Mark

; APPLICANT: Drake, Jim

; APPLICANT: Dimanac, Radoje

; APPLICANT: Engleman, Carrie

; APPLICANT: Faulkner, Brandy

; APPLICANT: Fox, Melvin

; APPLICANT: Garcia, Veronica

; APPLICANT: Giedt, Gretchen

; APPLICANT: Jomek, Leni

; APPLICANT: Jones, Lee

; APPLICANT: Kita, David

; APPLICANT: Kofler, Janette

; APPLICANT: Labat, Ivan

; APPLICANT: Lee, Won-Jae

; APPLICANT: Lomelli, Michelle

; APPLICANT: Nguyen, Hong

; APPLICANT: Nguyen, Linh

; APPLICANT: Nguyen, Lynne

; APPLICANT: Nguyen, Phuong

; APPLICANT: Nogra, Margie

; APPLICANT: Ojeda, Jesse

; APPLICANT: Palencia, Servando

; APPLICANT: Raisi, Fariba

; APPLICANT: Randhwa, Gurpreet

; APPLICANT: Sahourieh, Hannah

; APPLICANT: Sidhu, Navjivan

; APPLICANT: Smith, Benjamin

; APPLICANT: Smythe, Ashleigh

; APPLICANT: Tkach, Joe

; APPLICANT: Tulpule, Mukul

; APPLICANT: Verna, Ron

; APPLICANT: Wachter, Adam

; APPLICANT: Wu, James

; APPLICANT: Yim, Kenneth

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE REFERENCE: 754CIP

; CURRENT APPLICATION NUMBER: US/09/515,694

; CURRENT FILING DATE: 2000-02-29

; EARLIER APPLICATION NUMBER: 09/332,782

; EARLIER FILING DATE: 1999-06-14

; EARLIER APPLICATION NUMBER: 09/181,430

; EARLIER FILING DATE: 1998-10-28

; NUMBER OF SEQ ID NOS: 21027

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 8570

; LENGTH: 381

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

```
; NAME/KEY: misc_feature
; LOCATION: (1)...(381)
; OTHER INFORMATION: n = A,T,C or G
US-09-515-694-8570

Query Match      45.6%; Score 68; DB 57; Length 381;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 49 AGTGGATCCCGCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 97 agtggatccccggctgcaggaaattcgatcaagcttatcgataccgtcgacctcgag 156
    |||||||

QY 109 GGGGGGCC 116
    |||||||
Db 157 ggggggcc 164
```

```
RESULT 15
US-09-411-999-13627
; Sequence 13627, Application US/09411999
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Shenk, Michael Andrew
; TITLE OF INVENTION: Polynucleotides isolated from plants and
; FILE OF INVENTION: methods for their use.
; FILE REFERENCE: 1023U
; CURRENT APPLICATION NUMBER: US/09/411,999
; CURRENT FILING DATE: 1999-10-04
; NUMBER OF SEQ ID NOS: 37897
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13627
; LENGTH: 393
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-411-999-13627
```

```
Query Match      45.6%; Score 68; DB 54; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCGCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2 agtggatccccggctgcaggaaattcgatcaagcttatcgataccgtcgacctcgag 61
    |||||||

QY 109 GGGGGGCC 116
    |||||||
Db 62 ggggggcc 69
```

Search completed: May 30, 2000, 09:48:32
Job time: 60182 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:13:16 ; Search time 2192.43 Seconds
(without alignments)
275.462 Million cell updates/sec

Title: US-08-935-377-8
Perfect score: 149
Sequence: 1 GCCCAAAATGAAAACTA.....TTGTTTTGTGGGCCCCGGCC 149

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues 9714632

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 100000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	em_est1:*
2:	em_est2:*
3:	em_est3:*
4:	em_est4:*
5:	em_est5:*
6:	em_est6:*
7:	em_est7:*
8:	em_est8:*
9:	em_est9:*
10:	em_est10:*
11:	em_est11:*
12:	em_est12:*
13:	em_est13:*
14:	em_est14:*
15:	em_est15:*
16:	em_est16:*
17:	em_est17:*
18:	em_est18:*
19:	em_est19:*
20:	gb_est1:*
21:	gb_est2:*
22:	gb_est3:*
23:	gb_est4:*
24:	gb_est5:*
25:	gb_est6:*
26:	gb_est7:*
27:	gb_est8:*
28:	gb_est9:*
29:	gb_est10:*
30:	gb_est11:*
31:	gb_est12:*
32:	gb_est13:*
33:	gb_est14:*
34:	gb_est15:*
35:	gb_est16:*
36:	gb_est17:*
37:	gb_est18:*
38:	gb_est19:*
39:	gb_est20:*
40:	gb_est21:*
41:	gb_est22:*
42:	gb_est23:*
43:	gb_est24:*

44:	gb_est25:*
45:	gb_est26:*
46:	gb_est27:*
47:	gb_est28:*
48:	gb_est29:*
49:	em_est20:*
50:	em_est21:*
51:	em_est22:*
52:	gb_est30:*
53:	gb_est31:*
54:	gb_est32:*
55:	em_est23:*
56:	em_est24:*
57:	em_est25:*
58:	em_est26:*
59:	gb_est33:*
60:	gb_est34:*
61:	gb_est35:*
62:	gb_est36:*
63:	gb_est37:*
64:	gb_est38:*
65:	em_est27:*
66:	em_est28:*
67:	em_est29:*
68:	em_est30:*
69:	gb_est39:*
70:	gb_est40:*
71:	gb_est41:*
72:	gb_est42:*
73:	gb_est43:*
74:	gb_est44:*
75:	em_est31:*
76:	em_est32:*
77:	em_est33:*
78:	em_est34:*
79:	gb_est45:*
80:	gb_est46:*
81:	gb_est47:*
82:	gb_gss1:*
83:	gb_gss2:*
84:	gb_gss3:*
85:	gb_gss4:*
86:	em_gss1:*
87:	em_gss2:*
88:	em_gss3:*
89:	em_gss4:*
90:	gb_gss5:*
91:	gb_gss6:*
92:	em_gss5:*
93:	em_gss6:*
94:	gb_gss7:*
95:	gb_gss8:*
96:	gb_gss9:*
97:	em_gss7:*
98:	em_gss8:*
99:	gb_gss11:*
100:	gb_gss10:*
101:	em_gss9:*
102:	em_gss10:*
103:	em_gss11:*
104:	em_gss12:*
105:	gb_gss12:*
106:	gb_gss13:*
107:	gb_gss14:*
108:	gb_gss15:*
109:	gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
C 1	71	47.7	682	96	AQ074693 CIT-HSP-2
C 2	69	46.3	608	96	AQ009167 CIT-HFP-2
C 3	68.4	45.9	528	84	B69688 CIT-978SK-A
C 4	68	45.6	108	84	B54292 CIT-HSP-201
C 5	68	45.6	141	44	AU037147 AU037147
C 6	68	45.6	141	96	B95279 CIT-HSP-217
C 7	68	45.6	144	84	B80294 CIT-HSP-204
C 8	68	45.6	169	96	AQ013255 CIT-HSP-2
C 9	68	45.6	175	96	AQ041198 CIT-HSP-2
C 10	68	45.6	259	20	T48593 PH6f4_19/IT
C 11	68	45.6	275	27	AA052885 T3551 B10
C 12	68	45.6	386	84	B74162 CIT-HSP-202
C 13	68	45.6	462	96	AQ014512 CIT-HSP-2
C 14	68	45.6	519	84	B85534 CIT-HSP-202
C 15	68	45.6	547	96	AQ012629 CIT-HSP-2
C 16	68	45.6	552	96	AQ012216 CIT-HSP-2
C 17	68	45.6	575	96	AQ012221 CIT-HSP-2
C 18	68	45.6	596	96	AQ009165 CIT-HFP-2
C 19	68	45.6	607	84	B64363 CIT-HSP-202
C 20	68	45.6	632	84	B70299 CIT-HSP-202
C 21	68	45.6	638	96	AQ016737 CIT-HSP-2
C 22	68	45.6	647	84	B88775 CIT-HSP-205
C 23	68	45.6	666	84	B80296 CIT-HSP-204
C 24	68	45.6	700	96	AQ009153 CIT-HSP-2
C 25	67	45.0	541	96	AQ074695 CIT-HSP-2
C 26	67	45.0	576	38	AA752172 96S0243
C 27	66.4	44.6	513	96	AQ008437 CIT-HSP-2
C 28	66.4	44.6	541	96	AQ014519 CIT-HSP-2
C 29	63.8	42.8	134	84	B72088 CIT-978SK-13
C 30	63.2	42.4	233	96	AQ012365 CIT-HSP-2
C 31	63	42.3	234	20	T26395 AB161AAf IN
C 32	61.6	41.3	133	84	B26269 F10B5TF IGF
C 33	61.4	41.2	298	82	U68746 R1ftia pach
C 34	60	40.3	100	43	AI239372 SWOAFCAP
C 35	58.8	39.5	630	96	AQ003238 RPI111-19
C 36	57.8	38.8	145	84	B53695 CIT-HSP-201
C 37	57.4	38.5	603	38	AA752115 96S0066
C 38	57	38.3	144	84	B65391 CIT-HSP-202
C 39	56.8	38.1	97	41	AI058116 SWACAC15
C 40	56	37.6	471	96	AQ024018 CPGR0359A
C 41	56	37.6	499	96	AQ023564 CPGR0482A
C 42	56	37.6	547	96	AQ023548 CPGR0051A
C 43	56	37.6	626	96	AQ024077 CPGR0415A
C 44	55.4	37.2	502	96	AQ023674 CPGR0007B
C 45	55.2	37.0	149	96	AQ044029 CIT-HSP-2

ALIGNMENTS

RESULT 1
AQ074693/c
LOCUS CIT-HSP-2301L23.TF CIT-HSP Homo sapiens genomic clone 2301L23, genomic survey sequence.
DEFINITION genomic survey sequence.
ACCESSION AQ074693
VERSION AQ074693.1 GI:3436811
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 682)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Shizuya,H., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES

source
1..682
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2301L23"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII"
BASE COUNT 156 a 187 c 187 g 151 t 1 others
ORIGIN

Query Match 47.7%; Score 71; DB 96; Length 682;
Best Local Similarity 93.7%; Pred.No. 8.1e-14;
Matches 74; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 108
|||||
Db 79 AGTGGATCCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 20
|||||

QY 109 GGGGGCCCTAACTACTAA 127
|||||
Db 19 GGGGGCCCGAGTACCCAA 1
|||||

RESULT 2
AQ009167/c
LOCUS CIT-HFP-2281N13.TF CIT-HSP Homo sapiens genomic clone 2281N13, genomic survey sequence.
DEFINITION genomic survey sequence.
ACCESSION AQ009167
VERSION AQ009167.1 GI:3128599
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 608)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of a human BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other-GSSs: CIT-HSP-2281N13.FR.1 CIT-HSP-2281N13.TF
CIT-HSP-2281N13.FR

FEATURES

source
1..608

Best Local Similarity 98.6%; Pred. No. 5.9e-13;	
Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY 49	AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
Db 76	AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 17
QY 109	GGGGGGCCCTA 118
Db 16	GGGGGGCCCA 7
RESULT 4	
B54292	108 bp DNA GSS 20-JUN-1998
LOCUS	CIT-HSP-2017M1.TR CIT-HSP Homo sapiens genomic clone 2017M1,
DEFINITION	genomic survey sequence.
ACCESSION	B54292
VERSION	B54292.1 GI:2608626
KEYWORDS	GSS.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
REFERENCE	Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 108)
	Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
	Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
	Simon,M. and Venter,J.C.
TITLE	Use of a random BAC End Sequence Database for Sequence-Ready Map
JOURNAL	Building
COMMENT	Unpublished (1997)
	Contact: Mark Adams
	Department of Eukaryotic Genomics
	The Institute for Genomic Research
	9712 Medical Center Dr., Rockville, MD 20850, USA
	Tel: 301 838 0200
	Fax: 301 838 0208
	Email: mdadams@tigr.org
	Clones are available from Research Genetics (info@resgen.com). BAC
	end search page:
	http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
	Seq primer: M13, Reverse
	Class: BAC ends.
FEATURES	Location/Qualifiers
	1..108
	/organism="Homo sapiens"
	/db_xref="GDB:7044003"
	/db_xref="taxon:9606"
	/clone="2017M1"
	/clone_lib="CIT-HSP"
	/sex="Male"
	/cell_type="Sperm"
	/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
	HindIII"
BASE COUNT	20 a 35 c 34 g 19 t
ORIGIN	
	Query Match 45.6%; Score 68; DB 84; Length 108;
Best Local Similarity 100.0%; Pred. No. 6.3e-13;	
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 49	AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
Db 33	AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 92
QY 109	GGGGGCC 116
Db 93	GGGGGCC 100
RESULT 5	
AU037147	

LOCUS AU037147 141 bp mRNA EST 29-MAR-1999
 DEFINITION AU037147 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
 discoideum cDNA clone SSB532, mRNA sequence.
 ACCESSION AU037147
 VERSION AU037147.1 GI:3983900
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum.
 ORGANISM Dictyostelium discoideum
 Eukaryota; Dictyostellida; Dictyostellium.
 REFERENCE 1 (bases 1 to 141)
 AUTHORS Morio,T., Urushihara,H., Salto,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
 Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
 Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
 TITLE The Dictyostelium developmental cDNA project: generation and
 analysis of expressed sequence tags from the first-finger stage of
 development
 JOURNAL DNA Res. 5 (6), 335-340 (1998)
 MEDLINE 99156227
 COMMENT On Jan 19, 1998 this sequence version replaced gi:2153012.
 Contact: Hideko Urushihara
 Institute of Biological Sciences
 University of Tsukuba
 3-3-10 Ten-nodal, Tsukuba, Ibaraki 305, Japan
 Email: d402huesakura.cc.tsukuba.ac.jp
 PROJECT = 'Dictyostelium discoideum cDNA project in Japan'.
 FEATURES
 Location/Qualifiers
 1..141
 /organism="Dictyostelium discoideum"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="SSB532"
 /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
 /dev_stage="slug"
 /dev_stage="slug"
 BASE COUNT 26 a 40 C 42 g 28 t 5 others
 ORIGIN
 Query Match 45.6%; Score 68; DB 44; Length 141;
 Best Local Similarity 100.0%; Pred. No. 6.5e-13;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 49 AGTGGATCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 108
 |||||||
 DB 68 AGTGGATCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 127
 |||||||
 QY 109 GGGGGGCC 116
 |||||||
 DB 128 GGGGGGCC 135
 |||||||
 RESULT 6
 B95279/c 141 bp DNA GSS 26-JUN-1998
 LOCUS CIT-HSP-2172N2.FF CIT-HSP Homo sapiens genomic clone 2172N2,
 DEFINITION genomic survey sequence.
 ACCESSION B95279
 VERSION B95279.1 GI:2977616
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 141)
 AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
 Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
 Simon,M. and Venter,J.C.
 TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
 Building (1998)
 JOURNAL Unpublished (1998)
 COMMENT Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdaamsetigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
 Seq primer: M13 Reverse
 Class: BAC ends.
 Location/Qualifiers
 1..144
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /db_xref="GDB:7054558"

LOCUS AU037147 141 bp mRNA EST 29-MAR-1999
 DEFINITION AU037147 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
 discoideum cDNA clone SSB532, mRNA sequence.
 ACCESSION AU037147
 VERSION AU037147.1 GI:3983900
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum.
 ORGANISM Dictyostelium discoideum
 Eukaryota; Dictyostellida; Dictyostellium.
 REFERENCE 1 (bases 1 to 141)
 AUTHORS Morio,T., Urushihara,H., Salto,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
 Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
 Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
 TITLE The Dictyostelium developmental cDNA project: generation and
 analysis of expressed sequence tags from the first-finger stage of
 development
 JOURNAL DNA Res. 5 (6), 335-340 (1998)
 MEDLINE 99156227
 COMMENT On Jan 19, 1998 this sequence version replaced gi:2153012.
 Contact: Hideko Urushihara
 Institute of Biological Sciences
 University of Tsukuba
 3-3-10 Ten-nodal, Tsukuba, Ibaraki 305, Japan
 Email: d402huesakura.cc.tsukuba.ac.jp
 PROJECT = 'Dictyostelium discoideum cDNA project in Japan'.
 FEATURES
 Location/Qualifiers
 1..141
 /organism="Dictyostelium discoideum"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="SSB532"
 /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
 /dev_stage="slug"
 /dev_stage="slug"
 BASE COUNT 26 a 40 C 42 g 28 t 5 others
 ORIGIN
 Query Match 45.6%; Score 68; DB 44; Length 141;
 Best Local Similarity 100.0%; Pred. No. 6.5e-13;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 49 AGTGGATCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 108
 |||||||
 DB 68 AGTGGATCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAG 127
 |||||||
 QY 109 GGGGGGCC 116
 |||||||
 DB 128 GGGGGGCC 135
 |||||||
 RESULT 7
 B80294 144 bp DNA GSS 24-OCT-1998
 LOCUS CIT-HSP-2045D19.TR CIT-HSP Homo sapiens genomic clone 2045D19,
 DEFINITION genomic survey sequence.
 ACCESSION B80294
 VERSION B80294.1 GI:2867317
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 144)
 AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
 Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
 Simon,M. and Venter,J.C.
 TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
 Building (1998)
 JOURNAL Unpublished (1998)
 COMMENT Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdaamsetigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
 Seq primer: M13 Reverse
 Class: BAC ends.
 Location/Qualifiers
 1..144
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

```

/clone="2045D19"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/notes="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"
HindIII"
BASE COUNT      34 a      43 c      41 g      26 t
ORIGIN

Query Match      45.6%; Score 68; DB 84; Length 144;
Best Local Similarity 100.0%; Pred. No. 6.6e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
      |||||||
Db 69 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 128
      |||||||
QY 109 GGGGGGCC 116
      |||||||
Db 129 GGGGGGCC 136
      |||||||

RESULT 8
AQQ13255/c      169 bp      DNA      GSS      06-JUN-1998
LOCUS      CIT-HSP-2299C22.TF CIT-HSP Homo sapiens genomic clone 2299C22,
DEFINITION      genomic survey sequence.
ACCESSION      AQQ13255
VERSION      AQQ13255.1 GI:3185820
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 169)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL      Unpublished (1998)
COMMENT      Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES
source
1..169
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2045D19"
/cell_type="Sperm"
/sex="Male"
/cell_type="Sperm"
/notes="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"
HindIII"
BASE COUNT      32 a      47 c      48 g      42 t
ORIGIN

Query Match      45.6%; Score 68; DB 96; Length 169;
Best Local Similarity 100.0%; Pred. No. 6.7e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
      |||||||
Db 72 AGTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 131
      |||||||
QY 109 GGGGGGCC 116
      |||||||
Db 132 GGGGGGCC 139
      |||||||

RESULT 10
T48593      259 bp      mRNA
LOCUS      T48593
DEFINITION      ph6f4_19/1TV Outward Alu-primed hncDNA library Homo sapiens cDNA
clone ph6f4_19/1TV, mRNA sequence.

```

```

QY 49 ACTGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
      |||||||
Db 102 ACTGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 43
      |||||||
QY 109 GGGGGGCC 116
      |||||||
Db 42 GGGGGGCC 35
      |||||||

RESULT 9
AQQ41198      175 bp      DNA      GSS      14-JUL-1998
LOCUS      CIT-HSP-2335D21.TR CIT-HSP Homo sapiens genomic clone 2335D21,
DEFINITION      genomic survey sequence.
ACCESSION      AQQ41198
VERSION      AQQ41198.1 GI:3310469
KEYWORDS      GSS.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 175)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL      Unpublished (1998)
COMMENT      Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source
1..175
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2335D21"
/cell_type="Sperm"
/sex="Male"
/cell_type="Sperm"
/notes="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"
HindIII"
BASE COUNT      43 a      50 c      47 g      35 t
ORIGIN

Query Match      45.6%; Score 68; DB 96; Length 175;
Best Local Similarity 100.0%; Pred. No. 6.8e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
      |||||||
Db 72 AGTGGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 131
      |||||||
QY 109 GGGGGGCC 116
      |||||||
Db 132 GGGGGGCC 139
      |||||||

RESULT 10
T48593      259 bp      mRNA
LOCUS      T48593
DEFINITION      ph6f4_19/1TV Outward Alu-primed hncDNA library Homo sapiens cDNA
clone ph6f4_19/1TV, mRNA sequence.

```

```

ACCESSION T48593
VERSION T48593.1 GI:642793
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 259)
AUTHORS Vinogradova,T.V., Lebedev,Y.B., Kopantzev,E.P., Wagner,L.L., and
Volik,S.V., Ermolaeva,O.D., Lavrentyeva,I., Monastyrskaya,G.S. and
Sverdlov,E.D.
TITLE Outward Alu-primed hncDNA library
JOURNAL Unpublished (1995)
COMMENT Contact: Sverdlov ED
Structure and Function of Human Genes
Shemyakin Institute of Bioorganic Chemistry
16/10 Miklukho-Maklaya, Moscow, 117871, Russia
Tel: 70953306529
Fax: 70953306538
Email: sverd@humgen.siocb.msk.su.
FEATURES
Source Location/Qualifiers
1..259
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="phf4_19/11v"
/clone_lib="Outward Alu-primed hncDNA library"
/note="vector: pGEM-32; Site_1: EcoRI; Site_2: BamHI; The
library was constructed as described in [Obradovic, D.,
Borodin, A.M., Kopantzev, E.P., Wagner, L.L., Volik, S.V.,
Ermolaeva, O.D., Lebedev, Y.B., Monastyrskaya, G.S.,
Sverdlov, E.D. (1993) Bioorganicheskaya khimiya, 20,
919-930]. This protocol is based on nested primer strategy
using Alu- specific primers (ALN3 and TC-65) that direct
the hncDNA synthesis outward of Alu repeats."
BASE COUNT 62 a 66 c 72 g 59 t
ORIGIN
Query Match 45.6%; Score 68; DB 20; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 49 AGTCGATCCCCGGCGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
|||||
Db 55 AGTGATCCCCCGCGCGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 114
QY 109 GGGGGGCC 116
Db 115 GGGGGGCC 122
RESULT 11
AA052885
LOCUS t3551 Bloodstream form of serodeme ILrat1.1 Trypanosoma brucei
DEFINITION t3551 Bloodstream form of serodeme ILrat1.1 Trypanosoma brucei
ACCESSION AA052885
VERSION AA052885.1 GI:1543898
KEYWORDS EST.
SOURCE Trypanosoma brucei brucei.
ORGANISM Trypanosoma brucei brucei.
REFERENCE 1 (bases 1 to 275)
AUTHORS Osanya,A., Murphy,N.B. and Pelle,R.
TITLE Trypanosoma brucei cDNAs
JOURNAL Unpublished (1996)
COMMENT On Nov 29, 1993 this sequence version replaced gi:637578.
Contact: Osanya A
Unit 3
International Livestock Research Institute
Box 30709, Nairobi, Kenya
Tel: 254 2 630 743
Fax: 254 2 631 499
Email: a.osanya@cnet.com
Seq primer: M13 primer.
FEATURES
source Location/Qualifiers
1..275
/organism="Trypanosoma brucei brucei"
/db_xref="taxon:5702"
/clone_lib="Bloodstream form of serodeme ILrat1.1"
/note="cDNAs were generated from poly (A+) enriched mRNA
prepared from different developmental stages of T.b.brucei
by reverse transcription followed by PCR amplification
using mini-exon and oligo(dT) primers. The cDNA generated
were utilized in RADES-PCR coupled with differential
hybridisations to identify differentially expressed mRNA
transcripts. The products which showed to be
differentially expressed were cloned pGEM -T vector.
These differentially expressed mRNA transcripts were (are
being) sequenced to generate differentially expressed
sequence tags."
BASE COUNT 61 a 84 c 74 g 56 t
ORIGIN
Query Match 45.6%; Score 68; DB 27; Length 275;
Best Local Similarity 100.0%; Pred. No. 7.3e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 49 AGTGGATCCCCCGGCGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
|||||
Db 48 AGTGGATCCCCCGGCGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 107
QY 109 GGGGGGCC 116
Db 108 GGGGGGCC 115
RESULT 12
B74162/c
LOCUS B74162 386 bp DNA GSS 24-OCT-1998
DEFINITION CIT-HSP-2028N1.TF CIT-HSP Homo sapiens genomic clone 2028N1,
genomic survey sequence.
ACCESSION B74162
VERSION B74162.1 GI:2769849
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 386)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: CIT-HSP-2028N1.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21
Class: BAC ends.
FEATURES
source Location/Qualifiers
1..386
/organism="Homo sapiens"
/db_xref="GDB:7048251"

```



```

/db_xref="taxon:9606"
/clone="2028N1"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/notes="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT      81 a  98 c  109 g  98 t
ORIGIN

Query Match      45.6%; Score 68; DB 84; Length 386;
Best Local Similarity 100.0%; Pred. No. 7.7e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGATCCCCGGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
|||||
Db 83 AGTGATCCCCGGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 24
|||||
QY 109 GGGGGGCC 116
|||||
Db 23 GGGGGGCC 16

RESULT 13
AQU014512/c
LOCUS      AQU014512.1 462 bp DNA GSS 06-JUN-1998
DEFINITION CIT-HSP-2300D21.TF CIT-HSP Homo sapiens genomic clone 2300D21,
genomic survey sequence.
ACCESSION  AQU014512
VERSION     AQU014512.1 GI:3184839
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 462)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
BUILDING  (1998)
JOURNAL    Unpublished (1998)
COMMENT    Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..462
                     /organism="Homo sapiens"
                     /db_xref="GDB:7154425"
                     /db_xref="taxon:9606"
                     /clone="2300D21"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      89 a  127 c  131 g  115 t
ORIGIN

Query Match      45.6%; Score 68; DB 96; Length 462;
Best Local Similarity 100.0%; Pred. No. 7.9e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 49 AGTGATCCCCGGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
|||||
Db 100 AGTGATCCCCGGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 41
|||||
QY 109 GGGGGGCC 116
|||||
Db 40 GGGGGGCC 33

```

```

RESULT 14
B65534/c
LOCUS      B65534 519 bp DNA GSS 21-JUN-1998
DEFINITION CIT-HSP-2022D24.TF CIT-HSP Homo sapiens genomic clone 2022D24,
genomic survey sequence.
ACCESSION  B65534
VERSION     B65534.1 GI:2639512
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 519)
AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
BUILDING  (1997)
JOURNAL    Unpublished (1997)
COMMENT    Other GSSs: CIT-HSP-2022D24.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..519
                     /organism="Homo sapiens"
                     /db_xref="GDB:7045730"
                     /db_xref="taxon:9606"
                     /clone="2022D24"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      110 a  137 c  148 g  124 t
ORIGIN

```

```

Query Match      45.6%; Score 68; DB 84; Length 519;
Best Local Similarity 100.0%; Pred. No. 8e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGATCCCCGGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
|||||
Db 83 AGTGATCCCCGGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 24
|||||
QY 109 GGGGGGCC 116
|||||
Db 23 GGGGGGCC 16

RESULT 15
AQU012629
LOCUS      AQU012629 547 bp DNA GSS 06-JUN-1998

```

DEFINITION CIT-HSP-2299F1.TR CIT-HSP Homo sapiens genomic clone 2299F1,
genomic survey sequence.
ACCESSION AQ012629
VERSION AQ012629.1 GI:3185194
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamads@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source
1..547
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GBB:7154669"
/db_xref="taxon:9606"
/clone="2299F1"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT 154 a 139 c 125 g 129 t
ORIGIN

Query Match 45.6%; Score 68; DB 96; Length 547;
Best Local Similarity 100.0%; Pred. No. 8.1e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 AGTGGATCCCGCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 108
|||||
Db 74 AGTGGATCCCGCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 133
|||||

QY 109 GGGGGGCC 116
|||||
Db 134 GGGGGGCC 141

Search completed: May 29, 2000, 21:13:16
Job time: 36515 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:35:38 ; Search time 1214.87 Seconds
(without alignments)
940.489 Million cell updates/sec

Title: US-08-935-377-9
Perfect score: 150
Sequence: 1 GGCAAAATGAAAACTA.....TTGTTTTGTGGCCCGGCC 150

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 882769 seqs, 3808571567 residues 1765538
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 1000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_om.*
4: gb_ov.*
5: gb_pat.*
6: gb_ph.*
7: gb_pil.*
8: gb_p12.*
9: gb_p11.*
10: gb_p12.*
11: gb_p13.*
12: gb_ro.*
13: gb_sts.*
14: gb_sy.*
15: gb_un.*
16: gb_v1.*
17: em_fun.*
18: em_hum1.*
19: em_hum2.*
20: em_in.*
21: em_om.*
22: em_or.*
23: em_ov.*
24: em_pat.*
25: em_ph.*
26: em_pl.*
27: em_ro.*
28: em_sts.*
29: em_sy.*
30: em_un.*
31: em_v1.*
32: gb_htg1.*
33: gb_htg2.*
34: gb_in1.*
35: gb_in2.*
36: em_ba1.*
37: em_ba2.*
38: em_hum3.*
39: em_hum4.*
40: gb_pr4.*
41: gb_htg3.*
42: gb_htg4.*
43: gb_htg5.*

44: gb_htg6.*
45: gb_htg7.*
46: em_htg1.*
47: em_htg2.*
48: em_htg3.*
49: em_hum5.*
50: gb_pl3.*
51: gb_pr5.*
52: gb_htg8.*
53: gb_htg9.*
54: gb_htg10.*
55: gb_htg11.*
56: gb_htg12.*
57: gb_htg13.*
58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	69.8	46.5	793	5	AR060142	AR060142 Sequence
C 2	69	46.0	259	5	A44281	A44281 Sequence 9
C 3	69	46.0	2958	14	ARBLKSM	X52326 pBluescript
C 4	69	46.0	2958	14	ARBLKSP	X52331 pBluescript
C 5	69	46.0	2961	14	ARBL2KSM	X52329 pBluescript
C 6	69	46.0	2961	14	ARBL2KSP	X52327 pBluescript
C 7	69	46.0	2964	14	SYNBLKSMV	L08784 Bluescribe
C 8	69	46.0	2964	14	SYNBLKSPV	L08785 Bluescribe
C 9	69	46.0	3306	14	SYNPBEN66	D85525 Cloning vec
C 10	69	46.0	4144	14	XXU35131	U35131 Plasmid pBS
C 11	69	46.0	4267	14	PRS304	U3436 yeast integ
C 12	69	46.0	4289	14	XXU35136	U35136 Plasmid pBS
C 13	69	46.0	4373	14	PRS306	U3438 yeast integ
C 14	69	46.0	4443	14	PRS303	U3435 yeast integ
C 15	69	46.0	4670	14	ASAJ5326	AJ005326 pGAII(+)
C 16	69	46.0	4670	14	ASAJ5329	AJ005329 pGAII(-)
C 17	69	46.0	4707	14	XXU2374	U2374 Cloning vec
C 18	69	46.0	4768	14	XXU25061	U25061 Cloning vec
C 19	69	46.0	4783	14	PRS314	U3440 yeast centr
C 20	69	46.0	4887	14	PRS316	U3442 yeast centr
C 21	69	46.0	4950	14	XXU25060	U25060 Cloning vec
C 22	69	46.0	4967	14	PRS313	U3439 yeast centr
C 23	69	46.0	5144	14	CVU23751	U33751 Cloning vec
C 24	69	46.0	5187	14	U34887	U34887 yeast integ
C 25	69	46.0	5228	14	XXU25059	U25059 Cloning vec
C 26	69	46.0	5504	14	PRS305	U3437 yeast integ
C 27	69	46.0	5634	14	CVU14125	U14125 Cloning vec
C 28	69	46.0	6018	14	PRS315	U3441 yeast centr
C 29	69	46.0	6340	14	ASAJ5323	AJ005323 PCF1(-) K
C 30	69	46.0	9655	14	SYNPR8V	AD001331 Cloning v
C 31	68	45.3	2998	14	XXU43955	U43955 Expression
C 32	67.6	45.1	400	13	G16589	G16589 human sts S
C 33	67.4	44.9	3240	8	AF015771	AF015771 Magnaport
C 34	67	44.7	144	14	SYNPLKRA	M2847 Cloning vec
C 35	67	44.7	147	14	SYNPLKRB	M2848 Cloning vec
C 36	67	44.7	2958	14	ARBLSKM	X52324 pBluescript
C 37	67	44.7	2958	14	ARBLSKP	X52325 pBluescript
C 38	67	44.7	2961	14	ARBL2SKM	X52330 pBluescript
C 39	67	44.7	2961	14	ARBL2SKP	X52328 pBluescript
C 40	67	44.7	2964	14	SYNBLDKPV	L08787 Bluescribe
C 41	67	44.7	2964	14	SYNBLKSMV	L08786 Bluescribe
C 42	67	44.7	3062	14	CVU47947	U47947 Cloning vec
C 43	67	44.7	3228	14	CVE7829	AJ007829 Cloning v
C 44	67	44.7	3345	14	XXU35235	U35235 Plasmid pBS
C 45	67	44.7	3357	14	XXU35132	U35132 Plasmid pBS

ALIGNMENTS

```

RESULT 1
AR060142/c
LOCUS AR060142 793 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 121 from patent US 5840540.
ACCESSION AR060142
VERSION AR060142.1 GI:5986592
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 793)
AUTHORS St. George-Hyslop, P.H., Rommens, J.M. and Fraser, P.E.
TITLE Nucleic acids encoding presenilin II
JOURNAL Patent: US 5840540-A 121 24-NOV-1998;
FEATURES
Location/Qualifiers
source 1..793
BASE COUNT 162 a 242 c 146 g 229 t 14 others
ORIGIN

Query Match 46.5%; Score 69.8; DB 5; Length 793;
Best Local Similarity 82.5%; Pred. No. 1.1e-12;
Matches 80; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 39 GCGCGCCATGACGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 98
|||||
DB 130 GGAGCGCGGGGGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 71
|||||

QY 99 TCGACCTCGAGGGGGGCGCTAACTAACTAATTTGTT 135
|||||
DB 70 TCGACCTCGAGGGGGGCGCTAACTAACTAATTTGTT 34
|||||

RESULT 2
A44281/c
LOCUS A44281 259 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 9 from Patent WO9511300.
ACCESSION A44281
VERSION A44281.1 GI:2299115
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE
1 (bases 1 to 259)
AUTHORS Chandley, A.C., Kun, M., Sharkey, A.M., Hargreave, T.B. and Cooke, H.J.
TITLE AZOOSPERMIA IDENTIFICATION AND TREATMENT
JOURNAL Patent: WO 9511300-A 9 27-APR-1995;
COMMENT MEDICAL RES COUNCIL (GB)
FEATURES
Location/Qualifiers
source 1..259
/organism="unidentified"
/db_xref="taxon:32644"
/clone_lib="MOUSE GENOMIC LIBRARY"
/clone="M3.2"
BASE COUNT 67 a 58 c 63 g 69 t 2 others
ORIGIN

Query Match 46.0%; Score 69; DB 5; Length 259;
Best Local Similarity 88.2%; Pred. No. 1.7e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 51 GTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCCGCTCGAGG 110
|||||
DB 112 GTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCCGCTCGAGG 53
|||||

QY 111 GGGGGCCTAACTAACTAATTTGTT 135
|||||
DB 52 GGGGGCCTAACTAACTAATTTGTT 28
|||||

```

```

RESULT 3
ARBLKSP
LOCUS ARBLKSP 2958 bp DNA circular SYN 10-MAY-1995
DEFINITION pBluescript KS(-) vector DNA, phagemid excised from lambda ZAP.
ACCESSION X52326
VERSION X52326.1 GI:58064
KEYWORDS artificial sequence; cloning vector; expression vector; vector.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE
1 (bases 1 to 2958)
AUTHORS Thomas, E.A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE
2 (bases 1 to 2958)
AUTHORS Short, J.M., Fernandez, J.M., Sorge, J.A. and Huse, W.D.
TITLE Lambda ZAP: a bacteriophage lambda expression vector with in vivo
excision properties
JOURNAL Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE 88319944
REFERENCE
3 (bases 1 to 2958)
AUTHORS Altling-Wees, M.A. and Short, J.M.
TITLE pBluescript II: gene mapping vectors
JOURNAL Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE 90067967
FEATURES
Location/Qualifiers
source 1..2958
/organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 1..2958
/feature="phagemid pBluescript KS(-)"
BASE COUNT 708 a 754 c 731 g 765 t
ORIGIN

Query Match 46.0%; Score 69; DB 14; Length 2958;
Best Local Similarity 88.2%; Pred. No. 2.3e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 51 GTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCCGCTCGAGG 110
|||||
DB 687 GTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCCGCTCGAGG 746
|||||

QY 111 GGGGGCCTAACTAACTAATTTGTT 135
|||||
DB 747 GGGGGCCTAACTAACTAATTTGTT 771
|||||

RESULT 4
ARBLKSP
LOCUS ARBLKSP 2958 bp DNA circular SYN 11-JUN-1998
DEFINITION pBluescript KS(+) vector DNA, phagemid excised from lambda ZAP.
ACCESSION X52331
VERSION X52331.1 GI:58065
KEYWORDS artificial sequence; cloning vector; expression vector; vector.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE
1 (bases 1 to 2958)
AUTHORS Thomas, E.A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems,
11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE
2 (bases 1 to 2958)
AUTHORS Short, J.M., Fernandez, J.M., Sorge, J.A. and Huse, W.D.
TITLE Lambda ZAP: a bacteriophage lambda expression vector with in vivo
excision properties
JOURNAL Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE 88319944
REFERENCE
3 (bases 1 to 2958)

```

AUTHORS Alting-Mees,M.A. and Short,J.M.
TITLE pBluescript II: gene mapping vectors
JOURNAL Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE 90067967
REFERENCE 4 (bases 1 to 2958)
AUTHORS Lampe,D.J., Grant,T.E. and Robertson,H.M.
TITLE Factors affecting transposition of the Himarl mariner transposon in vitro
JOURNAL Genetics 149 (1), 179-187 (1998)
MEDLINE 98250682
FEATURES Location/Qualifiers
source 1..2958
/organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 1..2958
/note="phagemid pBluescript KS(+)"
BASE COUNT 749 a 734 c 751 g 724 t
ORIGIN

Query Match 46.0%; Score 69; DB 14; Length 2958;
Best Local Similarity 88.2%; Pred. No. 2.3e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 687 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746
|||||

Qy 111 GGGGCCCTAACTAACTAATTTGTT 135
|||||
Db 747 GGGGCCCGGTACCCAGCTTTGTT 771
|||||

RESULT 5
ARBL2KSM
LOCUS ARBL2KSM 2961 bp DNA circular SYN 10-MAY-1995
DEFINITION pBluescript II KS(-) vector DNA, phagemid excised from lambda ZAPII.
ACCESSION X52329.1 GI:58060
VERSION X52329
KEYWORDS artificial sequence; cloning vector; expression vector; vector.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 2961)
AUTHORS Thomas,E.A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems, 11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE 2 (bases 1 to 2961)
AUTHORS Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
TITLE Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties
JOURNAL Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE 88319944
REFERENCE 3 (bases 1 to 2961)
AUTHORS Alting-Mees,M.A. and Short,J.M.
TITLE pBluescript II: gene mapping vectors
JOURNAL Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE 90067967
FEATURES Location/Qualifiers
source 1..2961
/organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 1..2961
/note="phagemid pBluescriptII KS(-)"
BASE COUNT 706 a 758 c 735 g 762 t
ORIGIN

Query Match 46.0%; Score 69; DB 14; Length 2961;
Best Local Similarity 88.2%; Pred. No. 2.3e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 687 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746
|||||

Qy 111 GGGGCCCTAACTAACTAATTTGTT 135
|||||
Db 747 GGGGCCCGGTACCCAGCTTTGTT 771
|||||

RESULT 7
SYNBLKSMV
LOCUS SYNBLKSMV 2964 bp DNA circular SYN 26-JUL-1993
DEFINITION Bluescribe KS Minus cloning vector.
ACCESSION L08784
VERSION L08784.1 GI:310728
KEYWORDS Synthetic construct DNA.
SOURCE synthetic construct
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 2964)
AUTHORS Gilbert,W.
TITLE Obtained from VecBase 3.0

Qy 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 687 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746
|||||

Qy 111 GGGGCCCTAACTAACTAATTTGTT 135
|||||
Db 747 GGGGCCCGGTACCCAGCTTTGTT 771
|||||

RESULT 6
ARBL2KSP
LOCUS ARBL2KSP 2961 bp DNA circular SYN 10-MAY-1995
DEFINITION pBluescript II KS(+) vector DNA, phagemid excised from lambda ZAPII.
ACCESSION X52327
VERSION X52327.1 GI:58061
KEYWORDS artificial sequence; cloning vector; expression vector; vector.
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 2961)
AUTHORS Thomas,E.A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1990) Thomas E.A., Stratagene Clonin Systems, 11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE 2 (bases 1 to 2961)
AUTHORS Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
TITLE Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties
JOURNAL Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE 88319944
REFERENCE 3 (bases 1 to 2961)
AUTHORS Alting-Mees,M.A. and Short,J.M.
TITLE pBluescript II: gene mapping vectors
JOURNAL Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE 90067967
FEATURES Location/Qualifiers
source 1..2961
/organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 1..2961
/note="phagemid pBluescriptII KS(+)"
BASE COUNT 747 a 738 c 755 g 721 t
ORIGIN

Query Match 46.0%; Score 69; DB 14; Length 2961;
Best Local Similarity 88.2%; Pred. No. 2.3e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 687 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 746
|||||

Qy 111 GGGGCCCTAACTAACTAATTTGTT 135
|||||
Db 747 GGGGCCCGGTACCCAGCTTTGTT 771
|||||

RESULT 7
SYNBLKSMV
LOCUS SYNBLKSMV 2964 bp DNA circular SYN 26-JUL-1993
DEFINITION Bluescribe KS Minus cloning vector.
ACCESSION L08784
VERSION L08784.1 GI:310728
KEYWORDS Synthetic construct DNA.
SOURCE synthetic construct
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 2964)
AUTHORS Gilbert,W.
TITLE Obtained from VecBase 3.0

JOURNAL COMMENT Unpublished (1991) These data and their annotation were supplied to GenBank by Will Gilbert under the auspices of the GenBank Curator Program. Bluescribe KS Minus - Cloning vector #TYPE DNA CIRCULAR ENTRY Bluescribe KS Minus - Cloning vector TITLE Bluescribe KS Minus - Cloning vector DATE 28-JAN-1987 #sequence 02-FEB-1987 #sequence 04-MAR-1987 #sequence 03-APR-1987 ACCESSION V80077 SOURCE artificial REFERENCE #number 1 #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge J. #Journal Gene (1987) in press #citation Sequence data from Stratagene #comment Sequence correction according to Stratagene COMMENT Obtained from Stratagene on floppy disc. Revised 02-FEB-1987 by F. Pfeiffer: 1409/10 'AT' to 'TA' to match revised sequence of pBR322 Revised 4-MAR-1987 to match sequence of pUC19 on request of Stratagene Polylinker region revised 03-APR-1987 according to Stratagene COMMENT The stand shown corresponds to pUC19c. As in the published sequence of pUC19c, the M13mp19 lacZ region is on the complementary strand. This vector contains the fl origin so that the minus strand can be obtained upon fl superinfection. KEYWORDS CROSSREFERENCE #parent VecBase(3):BlueM13m #parent VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3, VecSource(3):bgalks, GenBank(50):PF1 #brother VecBase(3):BlueKSp, VecBase(3):BlueSkM PARENT Features of BlueSkM (2964 bp) residue source phage f1 3- 458 5488-5943 pUC19c 460- 624 236- 400 T7 promoter 626- 645 1- 20 BlueKs-polylinker 653- 760 108- 1 (c) T3 promoter 772- 791 20- 1 (c) pUC19 795-2964 448-2617 pUC19c Conflict (cfl) and Mutations (mut): none PARENT Features of BlueSkM (2964 bp) residue source phage f1 3- 458 5488-5943 pUC19c 460- 624 236- 400 T7 promoter 626- 645 1- 20 BlueKs-polylinker 653- 760 108- 1 (c) T3 promoter 772- 791 20- 1 (c) pUC19 795-1031 237- 1 (c) pUC19 1032-2964 2686- 754 (c) pUC19 Conflict (cfl) and Mutations (mut): none FEATURE 643 start of T7-RNA synthesis 77A (c) start of T3-RNA synthesis 1976-2764 789-1 (c) Ap-R; b-lactamase POLYLINKER KpnI-DraII-ApaI-XbaI-SalI-ClaI-HindIII-EcoRV-EcoRI-PstII-SmaI-BamHI-SpeI-XbaI-NotI-XmaIII-BstXI-SacII-SacI SELECTION #resistance Ap #indicator beta-galactosidase SUMMARY BlueSkM #length 2964 #checksum 1589. Location/Qualifiers FEATURES

Source 1. 2964 /organism="synthetic construct" /db_xref="taxon:32630" BASE COUNT 708 a 756 c 735 g 765 t ORIGIN

Query Match 46.0%; Score 69; DB 14; Length 2964; Best Local Similarity 88.2%; Pred. NO. 2.3e-12; Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGCTCGAGGAATTCGATATCAAGCTATCGATACCGTCGACCTCGAGG 110
|||||
Db 687 GTGGATCCCCGGGCTCGAGGAATTCGATATCAAGCTATCGATACCGTCGACCTCGAGG 746
QY 111 GGGGGCCTAACTAACTAATTTTGT 135
|||||
Db 747 GGGGGCCTAACTAACTAATTTTGT 771

RESULT 8
SYNBLKSPV 2964 bp DNA circular SYN 26-JUL-1993
LOCUS Bluescribe KS Plus cloning vector.
DEFINITION L08785
ACCESSION L08785.1 GI:310729
VERSION
KEYWORDS Synthetic construct DNA.
SOURCE synthetic construct
ORGANISM artificial sequence.
REFERENCE 1. (bases 1 to 2964)
AUTHORS Gilbert, W.
TITLE Obtained from VecBase 3.0
JOURNAL Unpublished (1991)
COMMENT These data and their annotation were supplied to GenBank by Will Gilbert under the auspices of the GenBank Curator Program. Bluescribe KS Plus - Cloning vector #TYPE DNA CIRCULAR ENTRY Bluescribe KS Plus - Cloning vector TITLE Bluescribe KS Plus - Cloning vector DATE 28-JAN-1987 #sequence 02-FEB-1987 #sequence 04-MAR-1987 #sequence 03-APR-1987 ACCESSION V80078 SOURCE artificial REFERENCE #number 1 #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge J. #Journal Gene (1987) in press #citation Sequence data from Stratagene #comment Sequence correction according to Stratagene COMMENT Obtained from Stratagene on floppy disc. Revised 02-FEB-1987 by F. Pfeiffer: 1409/10 'AT' to 'TA' to match revised sequence of pBR322 Revised 4-MAR-1987 to match sequence of pUC19 on request of Stratagene Polylinker region revised 03-APR-1987 according to Stratagene COMMENT The stand shown corresponds to pUC19c. As in the published sequence of pUC19c, the M13mp19 lacZ region is on the complementary strand. This vector contains the fl origin so that the plus strand can be obtained upon fl superinfection. KEYWORDS CROSSREFERENCE #parent VecBase(3):BlueM13p #parent VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3, VecSource(3):bgalks, GenBank(50):PF1 #brother

VecBase(3):BlueKSp, VecBase(3):BlueSkp

PARENT

Features of BlueKSp (2964 bp)

residue source
 3- 458 5943-5488 (c) phage f1
 460- 624 236- 400 pUC19c
 626- 645 1- 20 T7 promoter
 653- 760 108- 1 (c) BlueKSp-polylinker
 772- 791 20- 1 (c) T3 promoter
 795-2964 448-2617 pUC19c
 Conflict (cfl) and Mutations (mut): none

PARENT

Features of BlueKSp (2964 bp)

residue source
 3- 458 5943-5488 (c) phage f1
 460- 624 449- 285 (c) pUC19
 626- 645 1- 20 T7 promoter
 653- 760 108- 1 (c) BlueKSp-polylinker
 772- 791 20- 1 (c) T3 promoter
 795-1031 237- 1 (c) pUC19
 1032-2964 2686- 754 (c) pUC19
 Conflict (cfl) and Mutations (mut): none

FEATURE

643 start of T7-RNA synthesis
 774 (c) start of T3-RNA synthesis
 1976-2764 789-1 (c) Ap-R; b-lactamase

POLYLINKER

KpnI-DraII-ApaI-XhoI-SalI-ClaiI-HindIII-EcoRV-EcoRI-PstI-
 SmaI-BamHI-SpeI-XbaI-NotI-XmaII-BstXI-SacII-SacI SELECTION
 #resistance Ap

#indicator beta-galactosidase

SUMMARY BlueKSp #length 2964 #checksum 690.

FEATURES

source

1. .2964
 /organism="synthetic construct"
 /db_xref="taxon:32630" 723 t

BASE COUNT

ORIGIN

Query Match 46.0%; Score 69; DB 14; Length 2964;
 Best Local Similarity 88.2%; Pred. No. 2.3e-12;
 Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 Qy 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGCTCGAGG 110
 Db 687 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGCTCGAGG 746
 Qy 111 GGGGCCCTAACTAATTTTGT 135
 Db 747 GGGGCCCGGTACCCAGCTTTTGT 771

RESULT 9

SYNBPEN66/c

LOCUS SYNBPEN66 3306 bp DNA circular SYN 06-FEB-1999
 DEFINITION Cloning vector pBEN66 DNA for aminoglycoside 3'-phosphotransferase,
 beta-lactamase, complete cds.
 ACCESSION D85525
 VERSION D85525.1 GI:1345433
 KEYWORDS plasmid; aminoglycoside 3'-phosphotransferase; beta-lactamase.
 SOURCE Cloning vector pBEN66 (lab_host:E.coli) plasmid:pBEN66 DNA.
 ORGANISM Cloning vector pBEN66
 artificial sequence; vectors.
 REFERENCE 1 (bases 1 to 3306)
 AUTHORS Yamamoto.Y.

Direct Submission

TITLE Submitted (23-MAY-1996) to the DDBJ/EMBL/GenBank databases.
 JOURNAL Yoshihiro Yamamoto, Hyogo College of Medicine, Department of
 Genetics; Mukogawa-cho 1-1, Nishinomiya, Hyogo 663, Japan
 (tel:0798-45-6587, Fax:0798-40-7639)

REFERENCE

AUTHORS

Yamamoto.Y. and Furuyama,J.

One-step disruption by circular DNA in Escherichia coli

JOURNAL

Unpublished (1996)

FEATURES

source

1. .3306
 /organism="Cloning vector pBEN66"
 /plasmid="pBEN66"
 /db_xref="taxon:47800"
 /lab_host="E.coli"
 19. .36
 /note="T3 promoter"
 260. .1075
 /gene="kan"
 260. .1075
 /gene="kan"
 /note="derived from Tn903; kanamycin resistance gene"
 /citation=[2]
 /codon_start=1
 /transl_table=11
 /product="aminoglycoside 3'-phosphotransferase"
 /protein_id="BAAL2824.1"
 /db_xref="GI:1345434"

CDS

translation="MSHQRETSRPLNSMDADLYGKWARDNVGSGATYRLY
 GKPAPELFLKHGKSVANDVTDEMVLNLTETFMPLTKHFIRTPDDAWLLTAIP
 GKTAFOVLEYPDSGENIVDALAVFLRLHSIPVCNCPFNSDVRLEAQSRNNGL
 VDSDFDERNGWPVEQVWREMKLLFPSPSVVTHGDFSLNLIIFDEGLKICIDVG
 RVGIADRYQDLAILNWLNCGLFSPSLQRLFKYQIDNPMKNLQFHLMLDEFF"
 complement(2209. .3069)
 /gene="amp"
 complement(2209. .3069)
 /gene="amp"
 /note="ampicillin resistance gene"
 /citation=[1]
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase"
 /protein_id="BAAL2825.1"
 /db_xref="GI:1345435"
 /translation="MSIQHFRVALIPFAFCLPVFAHPETLVKVKDAEDOLGARVGY
 IELDLSNGKILESFRPEERPFMSKFKVLCGAVLSRIDAGQEQLRRIHYQNDLVE
 YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGPKELTAFHNHGDHVTPL
 DRWPELNEAIPNDERDTPVAMATTLRLKLTGELLTLASROQLIDWMEADKVGPL
 LRSALPAGWFIADKSGAGERSGIINALGPDGKPSRIVVIYITGSAQTMDERNRQIA
 EIGASLKKHW"

CDS

BASE COUNT 854 a 800 c 790 g 862 t
 ORIGIN

Query Match 46.0%; Score 69; DB 14; Length 3306;
 Best Local Similarity 88.2%; Pred. No. 2.3e-12;
 Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGCTCGAGG 110
 Db 121 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGCTCGAGG 52

Qy 111 GGGGCCCTAACTAATTTTGT 135

Db 61 GGGGCCCGGTACCCAGCTTTTGT 37

RESULT 10

XXU35131

LOCUS XXU35131 4144 bp DNA circular SYN 26-SEP-1995
 DEFINITION Plasmid pBSL159 cloning vector, complete sequence.
 ACCESSION U35131
 VERSION U35131.1 GI:984907
 KEYWORDS cloning vector pBSL159.
 SOURCE Cloning vector pBSL159
 ORGANISM artificial sequence; vectors.
 REFERENCE 1 (bases 1 to 4144)
 AUTHORS Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
 TITLE Improved antibiotic-resistance gene cassettes and omega elements

Wed May 31 10:04:58 2000

```

for Escherichia coli vector construction and in vitro
deletion/insertion mutagenesis
Gene 160 (1), 63-67 (1995)
95354958
REFERENCE
MEDLINE
AUTHORS
Hengen,P.N.
TITLE
Direct Submission
JOURNAL
Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical
Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
Location/Qualifiers
FEATURES
1. 4144
/organism="Cloning vector pBSL159"
/plasmid="pBSL159"
/db_xref="taxon:42704"
complement(1129..1989)
/EC_number="3.5.2.6"
/codon_start=1
/transl_table=11
/function="ampicillin resistance"
/product="beta-lactamase"
/protein_id="AAC53618.1"
/db_xref="GI:984908"
/translation="MSIOHFRVALIPFFAFLPVFAHPETLVKVKDAEDQLGARVGY
IEIDLNSGKILSEFPERPMWTFKVLICGAVLSRIDAGQEOGRIRHYSDNDLVE
YSPYKHKHLDGNTVEELCSAATMSDNLTAALLTTIGGPKELTAFLNNKDHVTRL
DRWPELNDIAIPNDERDTPVPAMATTLKLLTGLLTILASQQILIDWMDKVGAPL
LRSLAPAGFIADKSGAGERSGRIIATLALPGPKPSRIIVITTSQATMDERNQIA
EIGASLIKHW"
complement(join(3704..4144,1..219))
/codon_start=1
/transl_table=11
/product="chloramphenicol acetyltransferase"
/protein_id="AAC53619.1"
/db_xref="GI:984909"
/translation="MEKKITGYTVDISQWRKEHFEAFQSAQCTYNQTVOLDITAF
LKTIVKKNHRYFAFHILARLNNAHPERFAMKMGELVDSVHPCTVTFHEQETIF
SSLSVEYDDDFRFLHYSDVACYGENLAYFPKFIENFFVSANPWYSFTSFDLNV
ANMDFAPVETGKYTGQKVLMLPLAIQVHVAVCDGFHVGRLMNELOQYCDENQGG
A"
BASE COUNT 1084 a 1009 c 959 g 1092 t
ORIGIN

Query Match 46.0%; Score 69; DB 14; Length 4144;
Best Local Similarity 88.2%; Pred. No. 2.3e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
Db 2726 GTGGATCCCCGGCTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2785

QY 111 GGGGCGCTAACTAACTATTTGTT 135
Db 2786 GGGGCGCGGTACCAAGCTTTGTT 2810

RESULT 11
PR304 4267 bp DNA circular SYN 14-SEP-1995
LOCUS PR304
DEFINITION Yeast integrative vector PR304 with TRP1 marker, complete
sequence.
ACCESSION U03436
VERSION U03436.1 GI:416305
KEYWORDS Cloning vector PR304.
SOURCE Cloning vector PR304.
ORGANISM artificial sequence; vectors.
REFERENCE
1 (bases 1 to 4267)
AUTHORS Sikorski,R.S. and Hieter,P.
TITLE A system of shuttle vectors and yeast host strains designed for
efficient manipulation of DNA in Saccharomyces cerevisiae
JOURNAL Genetics 122 (1), 19-27 (1989)
MEDLINE 89276910

for Escherichia coli vector construction and in vitro
deletion/insertion mutagenesis
Gene 160 (1), 63-67 (1995)
95354958
REFERENCE
MEDLINE
AUTHORS
Hengen,P.N.
TITLE
Direct Submission
JOURNAL
Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical
Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
Location/Qualifiers
FEATURES
1. 4144
/organism="Cloning vector pBSL159"
/plasmid="pBSL159"
/db_xref="taxon:42704"
complement(1129..1989)
/EC_number="3.5.2.6"
/codon_start=1
/transl_table=11
/function="ampicillin resistance"
/product="beta-lactamase"
/protein_id="AAC53618.1"
/db_xref="GI:984908"
/translation="MSIOHFRVALIPFFAFLPVFAHPETLVKVKDAEDQLGARVGY
IEIDLNSGKILSEFPERPMWTFKVLICGAVLSRIDAGQEOGRIRHYSDNDLVE
YSPYKHKHLDGNTVEELCSAATMSDNLTAALLTTIGGPKELTAFLNNKDHVTRL
DRWPELNDIAIPNDERDTPVPAMATTLKLLTGLLTILASQQILIDWMDKVGAPL
LRSLAPAGFIADKSGAGERSGRIIATLALPGPKPSRIIVITTSQATMDERNQIA
EIGASLIKHW"
complement(join(3704..4144,1..219))
/codon_start=1
/transl_table=11
/product="chloramphenicol acetyltransferase"
/protein_id="AAC53619.1"
/db_xref="GI:984909"
/translation="MEKKITGYTVDISQWRKEHFEAFQSAQCTYNQTVOLDITAF
LKTIVKKNHRYFAFHILARLNNAHPERFAMKMGELVDSVHPCTVTFHEQETIF
SSLSVEYDDDFRFLHYSDVACYGENLAYFPKFIENFFVSANPWYSFTSFDLNV
ANMDFAPVETGKYTGQKVLMLPLAIQVHVAVCDGFHVGRLMNELOQYCDENQGG
A"
BASE COUNT 1084 a 1009 c 959 g 1092 t
ORIGIN

Query Match 46.0%; Score 69; DB 14; Length 4144;
Best Local Similarity 88.2%; Pred. No. 2.3e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
Db 2726 GTGGATCCCCGGCTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2785

QY 111 GGGGCGCTAACTAACTATTTGTT 135
Db 2786 GGGGCGCGGTACCAAGCTTTGTT 2810

RESULT 11
PR304 4267 bp DNA circular SYN 14-SEP-1995
LOCUS PR304
DEFINITION Yeast integrative vector PR304 with TRP1 marker, complete
sequence.
ACCESSION U03436
VERSION U03436.1 GI:416305
KEYWORDS Cloning vector PR304.
SOURCE Cloning vector PR304.
ORGANISM artificial sequence; vectors.
REFERENCE
1 (bases 1 to 4267)
AUTHORS Sikorski,R.S. and Hieter,P.
TITLE A system of shuttle vectors and yeast host strains designed for
efficient manipulation of DNA in Saccharomyces cerevisiae
JOURNAL Genetics 122 (1), 19-27 (1989)
MEDLINE 89276910

REFERENCE
2 (bases 1 to 4267)
AUTHORS Stillman,D.J.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
and Molecular Biology, University of Utah Medical Center, Salt Lake
City, UT 84132, USA
Location/Qualifiers
FEATURES
1. 4267
/organism="Cloning vector pBSL97"
/db_xref="taxon:31827"
BASE COUNT 1128 a 965 c 1092 g 1082 t
ORIGIN

Query Match 46.0%; Score 69; DB 14; Length 4267;
Best Local Similarity 88.2%; Pred. No. 2.3e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
Db 1921 GTGGATCCCCGGCTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 1980

QY 111 GGGGCGCTAACTAACTATTTGTT 135
Db 1981 GGGGCGCGGTACCAAGCTTTGTT 2005

RESULT 12
XXU35136 4289 bp DNA circular SYN 26-SEP-1995
LOCUS XXU35136
DEFINITION Plasmid pBSL97 cloning vector, complete sequence.
ACCESSION U35136
VERSION U35136.1 GI:984923
KEYWORDS cloning vector pBSL97.
ORGANISM Cloning vector pBSL97
artificial sequence; vectors.
REFERENCE
1 (bases 1 to 4289)
AUTHORS Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
TITLE Improved antibiotic-resistance gene cassettes and omega elements
for Escherichia coli vector construction and in vitro
deletion/insertion mutagenesis
JOURNAL Gene 160 (1), 63-67 (1995)
MEDLINE 95354958
REFERENCE
2 (bases 1 to 4289)
AUTHORS Hengen,P.N.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical
Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
Location/Qualifiers
FEATURES
1. 4289
/organism="Cloning vector pBSL97"
/plasmid="pBSL97"
/db_xref="taxon:42709"
complement(804..1598)
/codon_start=1
/transl_table=11
/product="neomycin phosphotransferase"
/protein_id="AAC53629.1"
/db_xref="GI:984924"
/translation="MIEQDGLHAGSPAAMVERLFGYDWAQOTICSDAAVFLSNAQGR
PVLVFKTDLGKALNELODEAARLSWLTATPCPDHQAHRERARTMEAGLVDDQDLDE
LSSHLAPAEKVISIMADAMRRLHTLDPAATPCPDHQAHRERARTMEAGLVDDQDLDE
EHQGLAPAEFLARLAKRMPDGEDLVVTHGDACLPLNIMVNGRFSFGIDCGRLGVADRY
QDIALATRDIAELGEGWADRFVLVLYGIAAPDSQRIAFYRLDDEFF"
complement(3301..4161)
/EC_number="3.5.2.6"
/codon_start=1
/transl_table=11
/function="ampicillin resistance"
/product="beta-lactamase"
/protein_id="AAC53630.1"
/db_xref="GI:984925"

```



```
/translation="MSIQHFRVALIPFFAARCLPVFAHPETLVKVKDAEDQLGARVGY
IELDNRSKILLESFPERFPMSTFKVLLCGAVLSRIDAGOEQLGRIRHYSQNDLVE
YSPTEKLTIDGMTVRELCSAAITMSDNTAANLLTITGGPKELTAFLHNMGDHVTRL
DRWPELNEATPNDERTTMPVAMATTILKLLTGELLTLASRQQLIDWEADKVGAPL
LRSLALPAGWFADKSGAGERSGRIIAALGPDGKPSRIWIYITGSGQATMDERNRQIA
EIGASLIKHW"
BASE COUNT 1024 a 1157 c 1111 g 997 t
ORIGIN

Query Match 46.0%; Score 69; DB 14; Length 4289;
Best Local Similarity 88.2%; Pred. No. 2.3e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 2015 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2074
|||||

QY 111 GGGGGCCTAACTAACTAATTTGTT 135
|||||
Db 2075 GGGGGCCCGGTACCCAGCTTTTGT 2099
|||||

RESULT 13
PRS306 PRS306 4373 bp DNA circular SYN 14-SEP-1995
LOCUS Yeast integrative vector PRS306 with URA3 marker, complete
DEFINITION
ACCESSION U03438
VERSION U03438.1 GI:416307
KEYWORDS Cloning vector PRS306.
SOURCE Cloning vector PRS306
ORGANISM artificial sequence; vectors.
REFERENCE 1 (bases 1 to 4373)
AUTHORS Sikorski, R.S. and Hieter, P.
TITLE A system of shuttle vectors and yeast host strains designed for
efficient manipulation of DNA in Saccharomyces cerevisiae
JOURNAL Genetics 122 (1), 19-27 (1989)
MEDLINE 89276910
AUTHORS Stillman, D.J.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
and Molecular Biology, University of Utah Medical Center, Salt Lake
City, UT 84132, USA
FEATURES
source
1. 4373
/organism="Cloning vector PRS306"
/db_xref="taxon:31829"
BASE COUNT 1181 a 987 c 1106 g 1099 t
ORIGIN

Query Match 46.0%; Score 69; DB 14; Length 4373;
Best Local Similarity 88.2%; Pred. No. 2.4e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 2027 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2086
|||||

QY 111 GGGGGCCTAACTAACTAATTTGTT 135
|||||
Db 2087 GGGGGCCCGGTACCCAGCTTTTGT 2111
|||||

RESULT 14
PRS303 PRS303 4443 bp DNA circular SYN 14-SEP-1995
LOCUS Yeast integrative vector PRS303 with HIS3 marker, complete
DEFINITION
ACCESSION U03435
```

```
U03435.1 GI:416304
KEYWORDS Cloning vector PRS303.
SOURCE Cloning vector PRS303
ORGANISM artificial sequence; vectors.
REFERENCE 1 (bases 1 to 4443)
AUTHORS Sikorski, R.S. and Hieter, P.
TITLE A system of shuttle vectors and yeast host strains designed for
efficient manipulation of DNA in Saccharomyces cerevisiae
JOURNAL Genetics 122 (1), 19-27 (1989)
MEDLINE 89276910
AUTHORS Stillman, D.J.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-1993) David J. Stillman, Dept. of Cellular, Viral
and Molecular Biology, University of Utah Medical Center, Salt Lake
City, UT 84132, USA
FEATURES
source
1. 4443
/organism="Cloning vector PRS303"
/db_xref="taxon:31826"
BASE COUNT 1149 a 1048 c 1109 g 1137 t
ORIGIN

Query Match 46.0%; Score 69; DB 14; Length 4443;
Best Local Similarity 88.2%; Pred. No. 2.4e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 2107 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 2166
|||||

QY 111 GGGGGCCTAACTAACTAATTTGTT 135
|||||
Db 2167 GGGGGCCCGGTACCCAGCTTTTGT 2191
|||||

RESULT 15
ASAJ5326/c ASAJ5326 4670 bp DNA circular SYN 08-FEB-1999
LOCUS PGAII(+) KS positive selection cloning vector glts gene.
DEFINITION
ACCESSION AJ005326
VERSION AJ005326.1 GI:4028984
KEYWORDS glts gene; glutamate permease.
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 4670)
AUTHORS Gal, J.
TITLE Direct Submission
JOURNAL Submitted (03-APR-1998) Gal J., Institute for Biotechnology, Bay
Zoltan Foundation for Applied Research, Szeged, Derkovits fasor 2.,
6726, HUNGARY
REFERENCE 2 (bases 1 to 4670)
AUTHORS Gal, J., Szekeres, S., Schnell, R., Pongor, S., Simonsits, A. and
Kalmán, M.
TITLE A positive selection cloning system based on the glts gene of
Escherichia coli
JOURNAL Anal. Biochem. 266 (2), 235-238 (1999)
MEDLINE 99107575
FEATURES
source
1. 4670
/organism="synthetic construct"
/db_xref="taxon:32630"
852..2231
/gene="glts"
852..2231
/gene="glts"
/codon_start=1
/transl_table=11
/product="glutamate permease"
/protein_id="CAA06473.1"
..CDS
```

wed May 31 10:04:58 2000

/db xref="GI:4028985"
/translation="MEHLDLTLVAATLTLGRKLVHVSFLKKYTIPEPVAGGLL
VALALLVLRKSMGEVNFDMSLRDPMLAFFATIGLNANIASLRAGRVGIFLIVV
GLLVQNAIGIGMSLGLDPLMLAGSITLGGHGTGAWSKLFIERYGFTNATEV
AMACATFGLVGLGIGPVARQITLTKGNKSWPGPPSRSTVSIISNSCSFGDPLV
LERPPRWSNSPYSESYARVLYKHSTTPNGIPDDQEVPTAFERPDVGRMITSLVLI
ETALLAICLTVCKIYAOLLAGTAFELPTFCVGLFVGVILSNGLSIMGFYRVERAS
VLGNVSLFLAMALGLKLWELASLALPMLAILLVQTIFFMALYALFTWRMMGNVD
AAVLAAGHCFCFGLGATPTAANNQATIERFGPSHMAFLVVPVWGAFFFIDIVNALVIKL
YLMPIFAG"

BASE COUNT 1040 a 1165 c 1231 g 1234 t
ORIGIN

Query Match 46.08; Score 69; DB 14; Length 4670;
Best Local Similarity 88.2%; Pred. No. 2.4e-12;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 51 GTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 1500 GTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 1441
|||||
QY 111 GGGGCGCTAACTAATTTGTT 135
|||||
Db 1440 GGGGCGCGGTACCGAGCTTTTGT 1416
|||||

Search completed: May 29, 2000, 21:35:42
Job time: 36874 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 21:58:33 ; Search time 1446.77 Seconds
(without alignments)
25.940 Million cell updates/sec

Title: US-08-935-377-9

Perfect score: 150
Sequence: 1 GGCACAAATGAAACATA.....TTGTTTTGTGGCCCGGCC 150

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	46.0	259	1 Q87664	Mouse azoospermia
2	69	46.0	501	1 T04866	Nucleotide analogu
3	67	44.7	651	1 X20513	Polynucleotide Seq
4	67	44.7	752	1 V31294	E. coli J96 pathog
5	67	44.7	1091	1 T75006	Human endogenous r
6	67	44.7	1122	1 T75005	Human endogenous r
7	67	44.7	3792	1 Q48463	Plasmid pg+host4 c
8	67	44.7	4226	1 T39485	Human steroidogene
9	67	44.7	5234	1 Q48464	Plasmid pg+host5 c
10	67	44.7	6722	1 Q48465	Plasmid pg+host6 c
11	67	44.7	10529	1 V09028	Maize fluory2 gene
12	67	44.7	12814	1 X24730	Swedish-FAD APP ta
13	67	44.7	15692	1 X24731	London-FAD APP tar
14	67	44.7	15692	1 X24732	Swedish/London-FAD
15	67	44.7	15701	1 X24733	Swedish-FAD APP13
16	66.8	44.5	545	1 V68808	Human endogenous r
17	65.8	43.9	6596	1 V57377	Maize female-prefe
18	65.4	43.6	2640	1 Q26664	bDAT. cDNA encodin
19	65.4	43.6	5356	1 T43794	Plasmid pRIPAT (r
20	65	43.3	84	1 X02800	E. coli biotin DNA
21	65	43.3	3465	1 X02814	DE19731274 Seq ID
22	65	43.3	3481	1 X02815	DE19731274 Seq ID
23	65	43.3	3794	1 X02812	DE19731274 Seq ID
24	65	43.3	3810	1 X02813	DE19731274 Seq ID
25	63.6	42.4	545	1 T75010	Human endogenous r
26	62.4	41.6	685	1 V10190	Stealth virus nucl
27	62.4	41.6	685	1 V12003	Stealth virus plas
28	59	39.3	4164	1 T04575	Plasmid pAT-1 sequ
29	59	39.3	4164	1 V22271	pAT-1 (pSD544). DN
30	59	39.3	4164	1 V69740	Nucleotide sequenc
31	58.2	38.8	1949	1 X05602	Nucleotide sequenc
32	58	38.7	5178	1 T49875	pTet-Splice. Nucle
33	57.8	38.5	3198	1 T92702	Candida Carhol gen

34 57.8 38.5 3198 1 T92869
35 56.4 37.6 459 1 V57250
36 56 37.3 78 1 X02799
37 56 37.3 2150 1 T91037
38 56 37.3 2150 1 X01271
39 56 37.3 2973 1 V64254
40 56 37.3 3956 1 V64258
41 56 37.3 4088 1 V64255
42 56 37.3 4102 1 V64257
43 56 37.3 4583 1 V64256
44 54.4 36.3 1094 1 X29498
45 54 36.0 10504 1 V79503

Candida Carhol gen
Clone #4 from muta
E. coli biotin DNA
Yeast checkpoint c
Yeast RAD17 coding
Plasmid PKS varian
Plasmid pPK13/14 D
Plasmid pPK5/6 DNA
Plasmid pPK9/10 DN
Plasmid pPK7/8 DNA
S. aureus MurF ORF
DNA sequence of ex

ALIGNMENTS

RESULT 1

Q87664/c
ID Q87664 standard; DNA: 259 BP.

AC Q87664;

DE 06-MAR-1996 (first entry)

DE Mouse azoospermia factor (AZF) gene partial clone M3.2.

KW Azoospermia factor; AZF; male infertility; YRRM gene;

KW Y-chromosome; RNA recognition motif; ds.

OS Mus sp.

FH Key Location/Qualifiers

FT misc_difference 174

FT /*tag= a
/note= "base n at position 174 is not identified
in the specification"

FT misc_difference 197

FT /*tag= b
/note= "base n at position 197 is not identified
in the specification"

PN W09511300-A2.

PD 27-APR-1995.

PF 24-OCT-1994; G03344.

PR 22-OCT-1993; GB-021857.

PR 07-JUL-1994; GB-013760.

PA (MED1-) MEDICAL RES COUNCIL.

PI Chandley AC, Cooke HJ, Hargreave TB, Kun M, Sharkey AM;

DR WPI; 95-170221/22.

PT Nucleic acid encoding the human azoospermia factor, and probes and

PT antibodies specific for the sequence and encoded polypeptide - may

PT be used in the clinical diagnosis of male infertility

PS Disclosure: Fig 5; 40pp; English.

CC A mouse gene partial sequence (Q87664) showing homology to the human

CC azoospermia factor (AZF) YRRM gene was obtd. from mouse genomic phage

CC clone M3.2 isolated from a genomic library screened with MYK (YRRM1)

CC cDNA (Q87655). M3.2, cloned in Lambda 2001, was mapped to the short arm

CC of the mouse Y-chromosome. A full cDNA sequence (Q87665) was also

CC obtd.

SQ Sequence 259 BP; 67 A; 58 C; 63 G; 69 T;

Query Match 46.0%; Score 69; DB 1; Length 259;

Best Local Similarity 88.2%; Pred. No. 8.5e-16;

Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGTGAGGAATTCGATATCAACCTATCGATACCGTCGACCTCGAGG 110

DB 112 GTGGATCCCCGGGTGAGGAATTCGATATCAACCTATCGATACCGTCGACCTCGAGG 53

QY 111 GGGGGCCTAACTAACTAATTTTGT 135

DB 52 GGGGGCCTAACTAACTAATTTTGT 28

RESULT 2

T04866
ID T04866 standard; DNA: 501 BP.

AC T04866;

DT 28-JAN-1996 (first entry)
DE Nucleotide analogue treated with calf intestinal alkaline phosphatase.
KW Nucleotide analogue; alkaline phosphatase; ss.
FH Synthetic.
OS Key
FT misc_feature 1..501
FT Location/Qualifiers
FT /tag= "std. IUPAC codes used"
FT /note= "std. IUPAC codes used"
PN WO9524505-A.
PD 14-SEP-1995.
PF 07-MAR-1995; G00490.
PR 08-MAR-1994; EP-301636.
PA (AMSH) AMERSHAM INT PLC.
PI Reeve MA, Robinson PS;
DR WPI; 95-328290/42.
PT Modification of residual fluorescence labelled nucleotide analogues - to
PT prevent migration in electrophoretic sequencing gel and interference with
PT base calling of DNA chains
PT Example; Fig 1: 18pp; English.
PS The invention concerns the modification of residual fluorescence
CC labelled nt analogues to prevent migration in electrophoretic
CC sequencing gel and interference with base calling of DNA chains. The
CC modification involves the use of a phosphatase enzyme to remove at
CC least one 5'-phosphate gp. The nts were prepared using ABI Amplitaq
CC dye-terminating kt. The template used was 1 microg. of M13 mp8. The
CC primer was M13 Universal primer. Samples were recovered. Some were
CC subjected to calf intestinal alkaline phosphatase digestion. This
CC includes T04866. T04867 was a control nt. which was not treated with
CC alkaline phosphatase. The example shows that alkaline phosphatase
CC treatment causes removal of the dye-terminator artefacts and allows
CC for accurate base calling with the ABI analysis software.
SQ Sequence 501 BP; 98 A; 136 C; 140 G; 121 T;

Query Match 46.0%; Score 69; DB 1; Length 501;
Best Local Similarity 88.2%; Pred. No. 1.4e-15;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGTGAGGATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 110
DB 57 GTGGATCCCCGGGTGAGGATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 116

QY 111 GGGGGCCTAACTAATTTGTT 135
DB 117 GGGGGCCCGGTACCCAGCTTTGTT 141

RESULT 3
X20513
ID X20513 standard; DNA; 651 BP.
AC X20513;
DT 05-MAY-1999 (first entry)
DE Polynucleotide sequence from the genome of Treponema pallidum.
KW Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW enzyme production; ds.
OS Treponema pallidum.
PN WO9859034-A2.
PD 30-DEC-1998.
PF 23-JUN-1998; U13041.
PR 24-JUN-1997; US-050667.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Fraser CM;
DR WPI; 99-081273/07.
PT New isolated Treponema pallidum nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention
PT and therapy of T. pallidum infections, particularly syphilis
PS Claim 1: Page 257; 1150pp; English.
CC X20500-21243 represent polynucleotide sequences from the genome of
CC Treponema pallidum. The sequences can be used for detection,
CC diagnosis, characterisation, prevention and therapy for T. pallidum
CC infections, particularly syphilis. They can also be used for detecting
CC diseases related to Borrelia infections in animals, and for the
CC production of biosynthetic products such as enzymes.

SQ Sequence 651 BP; 180 A; 167 C; 152 G; 149 T;

Query Match 44.7%; Score 67; DB 1; Length 651;
Best Local Similarity 100.0%; Pred. No. 6.2e-15;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGTGAGGATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 110
DB 78 GTGGATCCCCGGGTGAGGATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 137

QY 111 GGGGGCC 117
DB 138 GGGGGCC 144

RESULT 4
V31294/c
ID V31294 standard; DNA; 752 BP.
AC V31294;
DT 01-OCT-1998 (first entry)
DE E. coli J96 pathogenicity island contig #108.
KW PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pher;
KW PAI V; pher; vaccine; protective immune response; ds.
OS Escherichia coli.
PN WO9822575-A2.
PD 28-NOV-1998.
PF 21-NOV-1997; U21347.
PR 14-OCT-1997; US-061953.
PR 22-NOV-1996; US-031626.
PA (HUMA-) HUMAN GENOME SCI INC.
PA (UYWI-) UNIV WISCONSIN.
PI Choi GH, Dillon PJ, Welch RA;
DR WPI; 98-312461/27.
PT New isolated uropathogenic E. coli nucleotide sequences - used to
PT develop products for the detection of pathogenic E. coli and to
PT elicit an immune response to pathogenic E. coli
PS Claim 21: Page 221-222; 250pp; English.
CC This sequence represents a E. coli strain J96 contig containing
CC pathogenicity island (PAI) sequences, and represents a nucleic acid
CC molecule of the invention. PAIs are large fragments of DNA which comprise
CC pathogenicity determinants. The sequences of the invention are taken from
CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near pher)
CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at
CC approximately 94 min (at pher) on the E. coli chromosome and is
CC approximately 160 kb in size. Antibodies specific to the proteins encoded
CC by the PAI open reading frames of the invention can be used in kits to
CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit
CC a protective immune response in an animal to the uropathogenic E. coli
CC strain J96.

SQ Sequence 752 BP; 162 A; 213 C; 203 G; 172 T;

Query Match 44.7%; Score 67; DB 1; Length 752;
Best Local Similarity 100.0%; Pred. No. 6.5e-15;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGTGAGGATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 110
DB 72 GTGGATCCCCGGGTGAGGATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 13

QY 111 GGGGGCC 117
DB 12 GGGGGCC 6

RESULT 5
T75006/c
ID T75006 standard; DNA; 1091 BP.
AC T75006;
DT 06-OCT-1997 (first entry)
DE Human endogenous retroviral sequence 6.
KW Breast cancer; tumour; B18Ag1; prognosis; diagnosis; vaccine; ss.

OS Human retrovirus.
PN WO9725431-A1.
PD 17-JUL-1997.
PF 10-JAN-1997; U00398.
PR 10-JAN-1996; US-587329.
PA (CORI-) CORIXA CORP. JM;
PI Frudakis TN, Smith JM;
PT WPI: 97-384982/35.
DR Endogenous human tumour-associated retroviral element, B18Ag1 - used
PT for the prognosis, diagnosis and monitoring of human cancers,
PT especially breast cancer
PS Claim 10; Page 31-32; 74pp; English.
CC Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14
CC and 11-22 (75003-10) were obt'd. by screening human genomic
CC libraries using human breast tumour-associated retroviral element
CC B18Ag1 (see also T75002) as probe. These non-contiguous sequences
CC lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the
CC retrovirus genome (see also T75001). B18Ag1 and the other
CC retroviral sequences can be used in genetic vaccines and for the
CC prognosis, diagnosis and monitoring of human breast cancer.
SQ Sequence 1091 BP; 79 A; 350 C; 97 G; 248 T;

Query Match 44.7%; Score 67; DB 1; Length 1091;
Best Local Similarity 100.0%; Pred. No. 7.4e-15;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 110
DB 137 GTGGATCCCCGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 78
QY 111 GGGGGCC 117
DB 77 GGGGGCC 71

RESULT 6
T75005/c
ID T75005 standard; DNA; 1122 BP.
AC T75005;
DT 06-OCT-1997 (first entry)
DE Human endogenous retroviral sequence 3.
KW Breast cancer; tumour; B18Ag1; prognosis; diagnosis; vaccine; ss.
OS Human retrovirus.
PN WO9725431-A1.
PD 17-JUL-1997.
PF 10-JAN-1997; U00398.
PR 10-JAN-1996; US-587329.
PA (CORI-) CORIXA CORP. JM;
PI Frudakis TN, Smith JM;
PT WPI: 97-384982/35.
DR Endogenous human tumour-associated retroviral element, B18Ag1 - used
PT for the prognosis, diagnosis and monitoring of human cancers,
PT especially breast cancer
PS Claim 10; Page 30-31; 74pp; English.
CC Human endogenous retroviral sequences 10, 11-29, 3, 6, 12, 13, 14
CC and 11-22 (75003-10) were obt'd. by screening human genomic
CC libraries using human breast tumour-associated retroviral element
CC B18Ag1 (see also T75002) as probe. These non-contiguous sequences
CC lie in order 11-22, 14, B18Ag-1, 13, 12, 10, 3, 11-29, 6 in the
CC retrovirus genome (see also T75001). B18Ag1 and the other
CC retroviral sequences can be used in genetic vaccines and for the
CC prognosis, diagnosis and monitoring of human breast cancer.
SQ Sequence 1122 BP; 260 A; 316 C; 185 G; 279 T;

Query Match 44.7%; Score 67; DB 1; Length 1122;
Best Local Similarity 100.0%; Pred. No. 7.4e-15;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 110
DB 108 GTGGATCCCCGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 49

QY 111 GGGGGCC 117
DB 48 GGGGGCC 42

RESULT 7
Q48463/c
ID Q48463 standard; DNA; 3792 BP.
AC Q48463;
DT 18-MAR-1994 (first entry)
DE Plasmid pG+host4 containing Ts replication system.
KW Temperature sensitive replication; antibiotic resistance marker gene;
KW site-specific recombination; chromosomal integration; inactivation;
KW heterologous gene expression; thermosensitive plasmid; ds.
OS Synthetic.
PN WO9318164-A.
PD 16-SEP-1993.
PF 12-MAR-1993; F00248.
PR 13-MAR-1992; FR-003034.
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
PI Gruss A, Maguin E;
DR WPI: 93-303478/38.
PT New bacterial plasmid contg. heat sensitive replication system -
PT and marker gene, opt. capable of chromosomal integration, used to
PT inactivate specific gene or introduce heterologous gene
PS Example 2; Fig 9; 73pp; French.
CC Plasmid pG12 (Appl.Environ.Microbiol., 48: 726 (1984)) contg. two
CC antibiotic resistance marker genes was subjected to mutagenesis with
CC hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
CC sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
CC the 3340bp fragment lacking the Cm resistance gene was ligated to a
CC 445bp PvuII fragment of pBluescript SK+ containing a multicloning
CC site, T7 and T3 promoters and sequencing primer binding sites. The
CC resulting plasmid was designated pVE6004 (or pG+host4 - Q48463). It
CC is thermosensitive in all hosts tested, including E.coli, and must be
CC maintained at 26 deg.C. 1249 A; 552 C; 742 G; 1249 T;
SQ Sequence 3792 BP; 1249 A; 552 C; 742 G; 1249 T;

Query Match 44.7%; Score 67; DB 1; Length 3792;
Best Local Similarity 100.0%; Pred. No. 1.1e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 110
DB 3541 GTGGATCCCCGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 3482
QY 111 GGGGGCC 117
DB 3481 GGGGGCC 3475

RESULT 8
T39485
ID T39485 standard; DNA; 4226 BP.
AC T39485;
DT 21-MAY-1997 (first entry)
DE Human steroidogenesis acute regulatory protein genomic DNA.
KW Human; steroidogenesis; acute regulatory protein; hstAR; analysis;
KW mutation; detection; prenatal; genetic defect; congenital; protein;
KW lipid adrenal hyperplasia; treatment; prevention; gene;
KW replacement therapy; hypercholesterolaemia; ds.
OS Homo sapiens.
PN WO9629338-A1.
PD 26-SEP-1996.
PF 22-MAR-1996; U03396.
PR 23-MAR-1995; US-410540.
PA (REGC) UNIV CALIFORNIA.
PA (UYPE-) UNIV PENNSYLVANIA.
PI Lin D, Miller WL, Strauss JF;
DR WPI: 96-44330/44.
PT Isolated human steroidogenesis acute regulatory protein gene - used

for detection of mutation(s) of this gene that cause congenital lipid adrenal hyperplasia.
 Claim 1: Pages 23-25; 89pp; English.
 The present sequence encodes the human steroidogenesis acute regulatory protein (hSTAR). The hSTAR gene can be analysed for mutations to detect (e.g. prenatally) genetic defects associated with congenital lipid adrenal hyperplasia (CAH), or its transmission to children. CAH can be treated by protein or gene replacement therapy, which can also be used to prevent or treat hypercholesterolaemia.
 A human adrenal cortex cDNA library was screened with a mouse STAR probe to isolate a 1.6 kb insert, including an ORF for a 285 residue protein. When it was cloned into pSPORT and expressed in COS-1 cells cotransfected with pP450sc abp PDX, it increased the level of pregnenolone synthesis from cholesterol or 20-alpha-hydroxycholesterol.
 Sequence 4226 BP; 940 A; 1132 C; 1144 G; 984 T;
 SQ

Query Match 44.7%; Score 67; DB 1; Length 4226;
 Best Local Similarity 100.0%; Pred. No. 1.2e-14;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
 Db 4149 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4208
 QY 111 GGGGGCC 117
 Db 4209 GGGGGCC 4215

RESULT 9
 Q48464/c
 ID Q48464 standard; DNA; 5234 BP.
 AC Q48464;
 DT 18-MAR-1994 (first entry)
 DE Plasmid pG+host5 containing Ts replication system.
 KW Temperature sensitive replication; antibiotic resistance marker gene;
 KW site-specific recombination; chromosomal integration; inactivation;
 KW heterologous gene expression; thermosensitive plasmid; ds.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 36..1496
 FT FT /*tag= a
 FT FT /standard_name= ORI
 FT FT /note= "origin of replication from pBR322"
 FT FT 2640..4383
 FT FT /*tag= b
 FT FT /note= "from pGK12 (derived from pWV01)"
 FT FT 4384..4786
 FT FT /*tag= c
 FT FT /note= "from PUB110"
 FT FT 4787..5234
 FT FT /*tag= d
 FT FT /note= "from pSR"
 FN WO9318164-A.
 PD 16-SEP-1993.
 PF 12-MAR-1993; F00248.
 PR 13-MAR-1992; FR-003034.
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PI Gruss A, Maguin E;
 DR WPI; 93-303478/38.
 PT New bacterial plasmid contg. heat sensitive replication system -
 PT and marker gene, opt. capable of chromosomal integration, used to
 PT inactivate specific gene or introduce heterologous gene
 PS Disclosure: Fig 11; 73pp; French.
 CC Plasmid pGK12 (Appl. Environ. Microbiol., 48; 726 (1984)) contg. two
 CC antibiotic resistance marker genes was subjected to mutagenesis with
 CC hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
 CC sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
 CC the 3340bp fragment lacking the Cm resistance gene was ligated to a
 CC 445bp pvuII fragment of pBluescript SK+ containing a multicloning
 CC site. T7 and T3 promoters and sequencing primer binding sites. The
 CC resulting plasmid was designated pV56004 (or pG+host4 - Q48463).
 CC To facilitate cloning in E.coli, the Avai-EcoRI fragment of pBR322
 CC (containing the origin of replication and the ampicillin resistance
 CC gene) was inserted into NsiI-cleaved, blunt-ended pG+host4 to give
 CC pG+host6 (Q48466). The pBR322 ORI allows maintenance of the plasmid in
 CC E.coli at 37 deg.C. while the heat-sensitive ORI allows maintenance
 CC at 28 deg.C in gram-positive bacteria.
 SQ Sequence 6722 BP; 1961 A; 1278 C; 1503 G; 1980 T;

Query Match 44.7%; Score 67; DB 1; Length 6722;
 Best Local Similarity 100.0%; Pred. No. 1.4e-14;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
 Db 6471 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 6412
 QY 111 GGGGGCC 117
 Db 6411 GGGGGCC 6405

CC site, T7 and T3 promoters and sequencing primer binding sites. The
 CC resulting plasmid was designated pV56004 (or pG+host4 - Q48463).
 CC To facilitate cloning in E.coli, the 1.4kb Avai-AlwNI fragment of
 CC pBR322 (containing the origin of replication) was inserted into
 CC NsiI-cleaved pG+host4 to give pG+host5 (Q48464). The pBR322 ORI
 CC allows maintenance of the plasmid in E.coli at 37 deg.C. while the
 CC heat-sensitive ORI allows maintenance at 28 deg.C in gram-positive
 CC bacteria.
 SQ Sequence 5234 BP; 1569 A; 973 C; 1133 G; 1559 T;

Query Match 44.7%; Score 67; DB 1; Length 5234;
 Best Local Similarity 100.0%; Pred. No. 1.3e-14;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
 Db 4983 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4924

QY 111 GGGGGCC 117
 Db 4923 GGGGGCC 4917

RESULT 10
 Q48465/c
 ID Q48465 standard; DNA; 6722 BP.
 AC Q48465;
 DT 18-MAR-1994 (first entry)
 DE Plasmid pG+host6 containing Ts replication system.
 KW Temperature sensitive replication; antibiotic resistance marker gene;
 KW site-specific recombination; chromosomal integration; inactivation;
 KW heterologous gene expression; thermosensitive plasmid; ds.
 OS Synthetic.
 FN WO9318164-A.
 PD 16-SEP-1993.
 PF 12-MAR-1993; F00248.
 PR 13-MAR-1992; FR-003034.
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PI Gruss A, Maguin E;
 DR WPI; 93-303478/38.
 PT New bacterial plasmid contg. heat sensitive replication system -
 PT and marker gene, opt. capable of chromosomal integration, used to
 PT inactivate specific gene or introduce heterologous gene
 PS Disclosure: Fig 11; 73pp; French.
 CC Plasmid pGK12 (Appl. Environ. Microbiol., 48; 726 (1984)) contg. two
 CC antibiotic resistance marker genes was subjected to mutagenesis with
 CC hydroxylamine. A heat-stable mutant was isolated (coding for a heat-
 CC sensitive RepA - see Q48466 and Q48467), cut with ClaI and HpaII and
 CC the 3340bp fragment lacking the Cm resistance gene was ligated to a
 CC 445bp pvuII fragment of pBluescript SK+ containing a multicloning
 CC site. T7 and T3 promoters and sequencing primer binding sites. The
 CC resulting plasmid was designated pV56004 (or pG+host4 - Q48463).
 CC To facilitate cloning in E.coli, the Avai-EcoRI fragment of pBR322
 CC (containing the origin of replication and the ampicillin resistance
 CC gene) was inserted into NsiI-cleaved, blunt-ended pG+host4 to give
 CC pG+host6 (Q48466). The pBR322 ORI allows maintenance of the plasmid in
 CC E.coli at 37 deg.C. while the heat-sensitive ORI allows maintenance
 CC at 28 deg.C in gram-positive bacteria.
 SQ Sequence 6722 BP; 1961 A; 1278 C; 1503 G; 1980 T;

```

RESULT 11
V09028/c
ID V09028 standard; DNA; 10529 BP.
AC V09028.
DT 20-JUL-1998 (first entry)
DE Maize fluory2 gene (fl2).
KW Fluory2; fl2 gene; alpha-zeln; signal peptide; transgenic plant;
KW seed; essential amino acid; animal feed; maize; rice; wheat;
KW barley; millet; sorghum; ds.
OS Zea mays.
FH Key Location/Qualifiers
FT misc_feature 1..760
FT /tag= a
FT /note= "vector-derived sequence"
FT promoter 761..3824
FT /tag= b
FT /note= "Claim 14"
FT CDS 3825..4613
FT /tag= c
FT sig_peptide 3825..3887
FT /tag= d
FT mat_peptide 3888..4610
FT /tag= e
PN W09802563-Al.
PD 22-JAN-1998.
PF 11-JUL-1997; U11723.
PR 17-JUL-1996; US-021833.
PA (PION-) PIONEER HI-BRED INT INC.
PA (UYAR-) UNIV ARIZONA STATE.
PI Beach L, Coleman CE, Larkins BA;
DR WPI: 98-110609/10.
DR P-PSDB; W23977.
PT Cereal plants containing trans-gene expressing fusion that includes
PT signal peptide of the fl2 maize gene - and protein having high
PT content of essential amino acids, producing feeds of improved
PT nutritional value
PS Claim 14; Fig 1A-H; 37pp; English.
CC This is the nucleotide sequence of a clone of the fluory2 (fl2)
CC gene of maize. It codes for a 24-kDa alpha-zeln protein (see
CC W23977) that includes a 21-amino acid signal peptide (see W23976)
CC which targets the alpha-zeln to the lumen of the rough endoplasmic
CC reticulum. A claimed cereal plant contains a transgene comprising
CC a first polynucleotide that encodes the fl2 signal peptide and a
CC second polynucleotide that encodes an agronomically high-value
CC protein. Also new are seeds produced by the plants, the transgene
CC itself, and a transgene that also includes the fl2 promoter. The
CC second polynucleotide preferably encodes a protein that has a high
CC content of Met, Lys, Trp and/or Thr so that feeds from transformed
CC maize, wheat, rice, barley, millet or sorghum will have increased
CC contents of these essential amino acids in their seeds (all
CC claimed).
SQ Sequence 10529 BP; 2845 A; 2484 C; 2262 G; 2938 T;

Query Match 44.7%; Score 67; DB 1; Length 10529;
Best Local Similarity 100.0%; Pred. No. 1.6e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 110
|||||
DB 726 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 667
|||||
QY 111 GGGGGCC 117
DB 666 GGGGGCC 660

RESULT 12
X24730
ID X24730 standard; DNA; 12814 BP.
AC X24730;

```

```

DT 21-JUN-1999 (first entry)
DE Swedish-FAD APP targeting vector pMTI-2398.
KW Amyloid precursor protein; APP; human; gene targeting;
KW homologous recombination; transgenic mouse; transgenic animal;
KW animal model; Alzheimer's disease; vector; pMTI-2398;
KW Swedish-FAD APP; mutation; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT mat_peptide 1932..2276
FT /tag= a
FT /standard_name= Swedish-FAD APP
FT /note= "encodes W97997"
FT mat_peptide 5369..6160
FT /tag= b
FT /note= "encodes W97998"
PN W09909150-Al.
PD 25-FEB-1999.
PF 18-AUG-1997; U14507.
PR 18-AUG-1997; WO-U14507.
PA (FARB ) BAYER CORP.
PI Wirak DO;
DR WPI: 99-181029/15.
DR P-PSDB; W97997, W97998.
PT Modification of target nucleic acids - by homologous recombination,
PT used particularly for introducing a humanised amyloid precursor
PT protein gene into rodents for producing models of Alzheimer's
PT disease
PS Example; Page 96-103; 209pp; English.
CC This is the nucleotide sequence of Swedish-FAD APP targeting
CC vector pMTI-2398. The invention provides a novel gene targeting
CC strategy that facilitates the introduction of one or more specific
CC mutations into any gene in a single double reciprocal homologous
CC recombination step. The method has been used particularly for
CC introducing a humanised amyloid precursor protein (APP) gene into
CC rodents for producing animal models of Alzheimer's disease (AD).
CC 4 independent lines of transgenic mice (lines ES5007, ES5103,
CC ES5401 and ES5403) have been created using the gene targeting
CC technique applied to embryonic stem cells. In each line, the mouse
CC APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC amino acid residues 666-770 of APP770 were encoded by human CDNA
CC sequences instead of mouse genomic exons (exons 16-18). Within
CC these residues, only 3 amino acid differences exist between the
CC mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC Arg-684 to His. The exon-CDNA fusion gene therefore encodes an APP
CC containing a humanised beta-amyloid domain. In line ES5007,
CC created using vector pMTI-2398, the Swedish mutation, i.e.
CC KM(670,681)NL, was also introduced. The targeted Swedish-FAD
CC m/hAPP gene expressed m/hAPP protein at levels approaching those
CC observed for mouse APP in brain.
SQ Sequence 12814 BP; 1865 A; 1850 C; 1962 G; 1939 T;

Query Match 44.7%; Score 67; DB 1; Length 12814;
Best Local Similarity 100.0%; Pred. No. 1.7e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 110
|||||
DB 9888 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 9947
|||||
QY 111 GGGGGCC 117
DB 9948 GGGGGCC 9954

RESULT 13
X24731
ID X24731 standard; DNA; 15692 BP.
AC X24731.
DT 21-JUN-1999 (first entry)
DE London-FAD APP targeting vector pMTI-2453.
KW Amyloid precursor protein; APP; human; gene targeting;

```

homologous recombination; transgenic mouse; transgenic animal;
animal model; Alzheimer's disease; vector; pMTI-2453;
London-FAD APP; mutation; ss.
Synthetic.
Key Location/Qualifiers
mat_peptide 4807..5151
FT /*tag= a
FT /standard_name= "Swedish-FAD"
FT /note= "encodes W97999"
FT /*tag= b
FT /standard_name= "London-FAD"
FT 8223..9023
FT /*tag= C
FT /note= "encodes W97998"
W09909150-A1.
PN 25-FEB-1999.
PD 18-AUG-1997; U14507.
PF 18-AUG-1997; WO-U14507.
PR (FARB) BAYER CORP.
PA Wirak DO;
PI WPI; 99-181029/15.
DR P-PSDB; W97998, W97999.
DR Modification of target nucleic acids - by homologous recombination,
used particularly for introducing a humanised amyloid precursor
protein gene into rodents for producing models of Alzheimer's
disease
PT Example; Page 104-113; 209pp; English.
PS This is the nucleotide sequence of London-FAD APP targeting
vector pMTI-2453. The invention provides a novel gene targeting
strategy that facilitates the introduction of one or more specific
mutations into any gene in a single double reciprocal homologous
recombination step. The method has been used particularly for
introducing a humanised amyloid precursor protein (APP) gene into
rodents for producing animal models of Alzheimer's disease (AD).
CC 4 Independent lines of transgenic mice (lines ES5007, ES5103,
ES5401 and ES5403) have been created using the gene targeting
technique applied to embryonic stem cells. In each line, the mouse
APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
amino acid residues 666-770 of APP770 were encoded by human CDNA
sequences instead of mouse genomic exons (exons 16-18). Within
these residues, only 3 amino acid differences exist between the
mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
containing a humanised beta-amyloid domain. In line ES5401,
created using vector pMTI-2453, the London mutation, i.e.
V(717)L, was also introduced.
CC Sequence 15692 BP; 2423 A; 2730 C; 2786 G; 2541 T;
SQ Query Match 44.7%; Score 67; DB 1; Length 15692;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 110
Db 12766 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 12825
QY 111 GGGGGCC 117
Db 12826 GGGGGCC 12832
RESULT 14
X24732
ID X24732 standard; DNA; 15692 BP.
AC X24732;
DT 21-JUN-1999 (first entry)
DE Swedish/London-FAD APP targeting vector pMTI-2454.
DE Amyloid precursor protein; APP; human; gene targeting;
KW homologous recombination; transgenic mouse; transgenic animal;
KW animal model; Alzheimer's disease; vector; pMTI-2454;

London-FAD APP; Swedish-FAD APP; mutation; ss.
Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT mat_peptide 4807..5151
FT /*tag= a
FT /note= "encodes W98000"
FT replace(4849,"")
FT /*tag= b
FT /standard_name= "Swedish-FAD"
FT replace(4989,"")
FT /standard_name= "London-FAD"
FT 8223..9023
FT /*tag= b
FT /note= "encodes W97998"
W09909150-A1.
PN 25-FEB-1999.
PD 18-AUG-1997; U14507.
PF 18-AUG-1997; WO-U14507.
PR (FARB) BAYER CORP.
PA Wirak DO;
PI WPI; 99-181029/15.
DR P-PSDB; W97998, W98000.
DR Modification of target nucleic acids - by homologous recombination,
used particularly for introducing a humanised amyloid precursor
protein gene into rodents for producing models of Alzheimer's
disease
PT Example; Page 114-123; 209pp; English.
PS This is the nucleotide sequence of Swedish/London-FAD APP targeting
vector pMTI-2454. The invention provides a novel gene targeting
strategy that facilitates the introduction of one or more specific
mutations into any gene in a single double reciprocal homologous
recombination step. The method has been used particularly for
introducing a humanised amyloid precursor protein (APP) gene into
rodents for producing animal models of Alzheimer's disease (AD).
CC 4 Independent lines of transgenic mice (lines ES5007, ES5103,
ES5401 and ES5403) have been created using the gene targeting
technique applied to embryonic stem cells. In each line, the mouse
APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
amino acid residues 666-770 of APP770 were encoded by human CDNA
sequences instead of mouse genomic exons (exons 16-18). Within
these residues, only 3 amino acid differences exist between the
mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
containing a humanised beta-amyloid domain. In line ES5103,
created using vector pMTI-2454, the London mutation, i.e.
V(717)L, and the Swedish mutation, i.e. KM(670,671)NL, were
introduced. The targeted Swedish/London-FAD m/hAPP gene expressed
m/hAPP protein at levels approaching those observed for mouse APP
in brain.
CC Sequence 15692 BP; 2423 A; 2731 C; 2783 G; 2542 T;
SQ Query Match 44.7%; Score 67; DB 1; Length 15692;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 110
Db 12766 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTGCACCTCGAGG 12825
QY 111 GGGGGCC 117
Db 12826 GGGGGCC 12832
RESULT 15
X24733
ID X24733 standard; DNA; 15701 BP.
AC X24733;
DT 21-JUN-1999 (first entry)
DE Swedish-FAD APP713 targeting vector pMTI-2454.
DE Amyloid precursor protein; APP; human; gene targeting;
KW Amyloid precursor protein; APP; human; gene targeting;

KW homologous recombination; transgenic mouse; transgenic animal;
KW animal model; Alzheimer's disease; vector; pMTI-2455;
KW Swedish-FAD APP713; mutation; ss.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT mat_peptide 4807..4983
FT FT /*tag= a
FT FT /*note= "encodes W98001"
FT mutation replace(4835,"")
FT FT /*tag= b
FT FT /*standard_name= "Swedish-FAD"
FT mutation replace(4981,"")
FT FT /*tag= c
FT FT /*standard_name= APP713stop
FT mat_peptide 8232..9032
FT FT /*tag= d
FT FT /*note= "encodes W97998"
PN W09909150-A1.
PD 25-FEB-1999.
PF 18-AUG-1997; U14507.
PA 18-AUG-1997; WO-U14507.
PA (FARB) BAYER CORP.
PI Wirak DO;
DR WPI: 99-181029/15.
DR P-PSDB; W97998, W98001.
PT Modification of target nucleic acids - by homologous recombination,
PT used particularly for introducing a humanised amyloid precursor
PT protein gene into rodents for producing models of Alzheimer's
PT disease
PS Example; Page 124-133; 209pp; English.
CC This is the nucleotide sequence of Swedish-FAD APP713 targetting
CC vector pMTI-2455. The invention provides a novel gene targetting
CC strategy that facilitates the introduction of one or more specific
CC mutations into any gene in a single double reciprocal homologous
CC recombination step. The method has been used particularly for
CC introducing a humanised amyloid precursor protein (APP) gene into
CC rodents for producing animal models of Alzheimer's disease (AD).
CC 4 independent lines of transgenic mice (lines ES5007, ES5103,
CC ES5401 and ES5403) have been created using the gene targetting
CC technique applied to embryonic stem cells. In each line, the mouse
CC APP gene was modified to encode a mouse/human hybrid (m/hAPP) where
CC amino acid residues 666-770 of APP770 were encoded by human cDNA
CC sequences instead of mouse genomic exons (exons 16-18). Within
CC these residues, only 3 amino acid differences exist between the
CC mouse and human proteins, i.e. Gly-676 to Arg, Phe-681 to Thr and
CC Arg-684 to His. The exon-cDNA fusion gene therefore encodes an APP
CC containing a humanised beta-amyloid domain. In line ES5215,
CC created using vector pMTI-2455, the Swedish mutation. i.e.
CC KM(670,681)NL, was also introduced, and also a stop codon (T to
CC stop at position 714).
SQ Sequence 15701 BP; 2470 A; 2675 C; 2845 G; 2497 T;

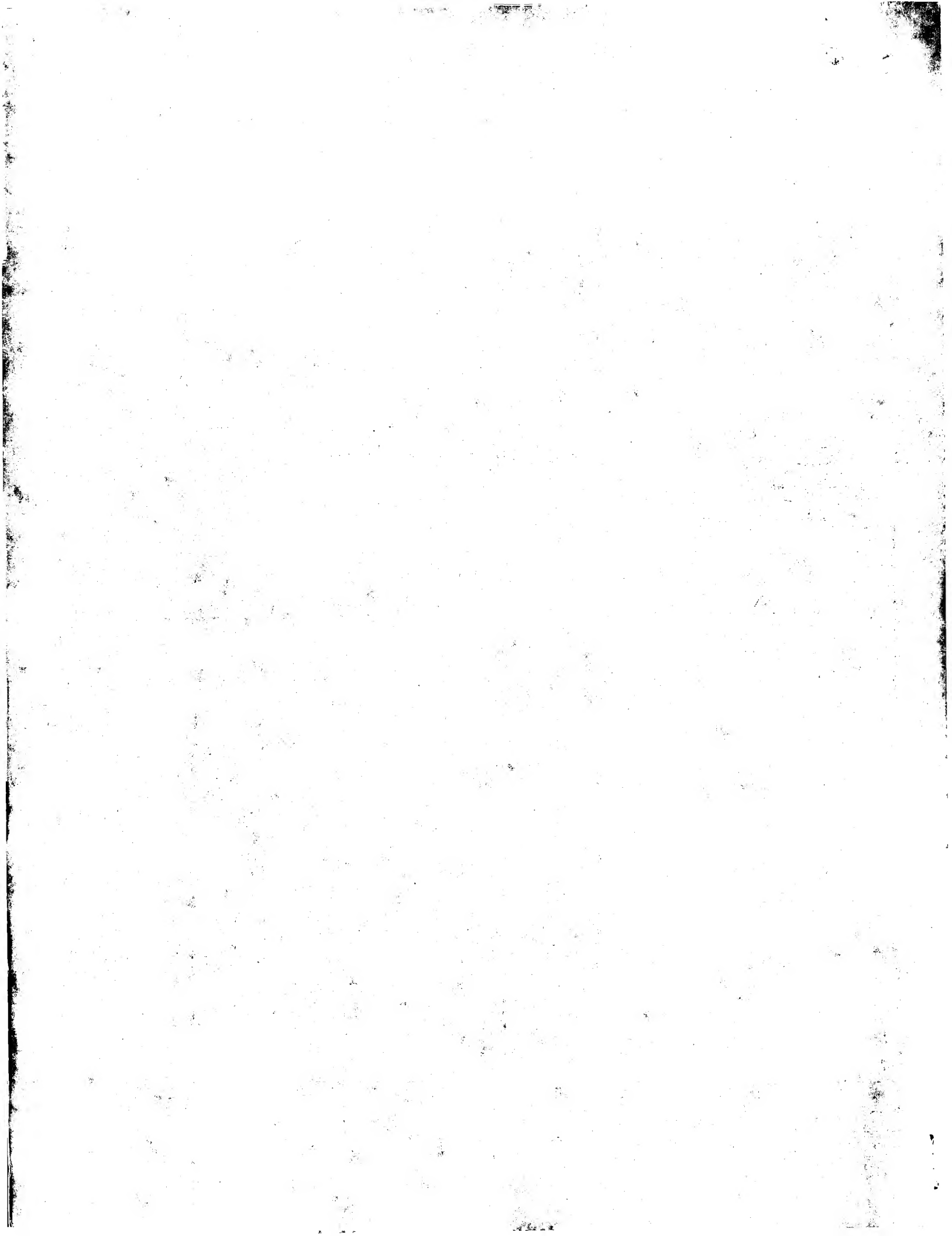
Query Match 44.7%; Score 67; DB 1; Length 15701;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGTGCAGGAATTCGATATCAAGCTATCGATACCGTCGACCTCGAGG 110
|||||
Db 12775 GTGGATCCCCGGGTGCAGGAATTCGATATCAAGCTATCGATACCGTCGACCTCGAGG 12834
|||||

QY 111 GGGGGCC 117
|||||

Db 12835 GGGGGCC 12841

Search completed: May 29, 2000, 21:58:40
Job time: 38188 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2000, 22:09:10 ; Search time 621.83 Seconds
(without alignments)
31.355 Million cell updates/sec

Title: US-08-935-377-9
Perfect score: 150
Sequence: 1 GCCCAAAATGAAAACTA.....TTGTTTTTGGCGCCGGCC 150

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgnl_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgnl_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgnl_6/ptodata/1/ina/5C.COMB.seq.*
4: /cgnl_6/ptodata/1/ina/5D.COMB.seq.*
5: /cgnl_6/ptodata/1/ina/6.COMB.seq.*
6: /cgnl_6/ptodata/1/ina/PTUS.COMB.seq.*
7: /cgnl_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	69.8	46.5	793	3	US-08-967-101-121
C 2	69.8	46.5	793	4	US-08-967-101-121
C 3	67	44.7	3792	4	US-08-992-341-121
C 4	67	44.7	3792	5	US-08-302-752-1
C 5	67	44.7	4016	2	US-08-410-540-3
C 6	67	44.7	5234	4	US-08-992-334-2
C 7	67	44.7	5234	5	US-08-302-752-2
C 8	67	44.7	6722	4	US-08-992-334-3
C 9	67	44.7	6722	5	US-08-302-752-3
C 10	66.2	44.1	9318	3	US-08-793-610-6
C 11	62.4	41.6	685	2	US-08-463-115-56
C 12	62.4	41.6	685	2	US-08-465-388-56
C 13	59	39.3	4164	1	US-08-204-675-1
C 14	59	39.3	4164	3	US-08-680-754-1
C 15	59	39.3	4164	4	US-08-796-364-1
C 16	59	39.3	4164	6	PCT-US95-02520-1
C 17	58	38.7	5178	3	US-08-474-169-2
C 18	56	37.3	88	1	US-08-144-602B-15
C 19	56	37.3	1023	1	US-08-198-446B-1
C 20	56	37.3	1023	3	US-08-870-693-1
C 21	56	37.3	2150	1	US-08-198-446B-10
C 22	56	37.3	2150	3	US-08-870-693-10
C 23	55	36.7	834	3	US-08-967-101-113
C 24	55	36.7	834	4	US-08-592-541-113
C 25	51.8	34.5	2308	1	US-08-325-071-62
C 26	51.2	34.1	3341	3	US-08-868-577-18

C 27	50	33.3	780	3	US-08-967-101-116	Sequence 116, App
C 28	50	33.3	780	4	US-08-992-341-116	Sequence 116, App
C 29	50	33.3	6505	3	US-08-793-610-5	Sequence 5, Appli
C 30	49.8	33.2	1200	1	US-08-011-398B-3	Sequence 3, Appli
C 31	49.8	33.2	1200	2	US-08-464-051-3	Sequence 3, Appli
C 32	49.8	33.2	1200	3	US-08-462-498-3	Sequence 3, Appli
C 33	49.8	33.2	1200	3	US-08-462-498-3	Sequence 3, Appli
C 34	49.8	33.2	6244	1	US-08-076-726-15	Sequence 15, Appli
C 35	49.8	33.2	6244	1	US-08-260-452-8	Sequence 8, Appli
C 36	49.8	33.2	6244	3	US-08-481-970-8	Sequence 8, Appli
C 37	48	32.0	1612	1	US-08-897-719-8	Sequence 8, Appli
C 38	48	32.0	2764	4	US-08-343-733A-1	Sequence 1, Appli
C 39	47.2	31.5	2185	4	US-08-465-971B-1	Sequence 1, Appli
C 40	47	31.3	1997	4	US-08-467-848A-3	Sequence 3, Appli
C 41	47	31.3	2126	3	US-08-667-809B-3	Sequence 3, Appli
C 42	47	31.3	2605	3	US-08-789-354-1	Sequence 1, Appli
C 43	47	31.3	5534	2	US-08-680-395-4	Sequence 4, Appli
C 44	46.6	31.1	1249	2	US-08-452-267-3	Sequence 3, Appli
C 45	46.6	31.1	1249	2	US-08-463-115-35	Sequence 35, Appli
C 45	46.6	31.1	1249	2	US-08-455-388-35	Sequence 35, Appli

ALIGNMENTS

RESULT 1
US-08-967-101-121/c
; Sequence 121, Application US/08967101
; Patent No. 584054C
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,101
; FILING DATE: 10-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 793 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-967-101-121

Query Match 46.5% Score 69.8; DB 3; Length 793;
Best Local Similarity 82.5%; Pred. No. 4.5e-16;
Matches 80; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

COMPUTER READABLE FORM:

REGISTRATION NUMBER: 20958

**A
C
C
E
P
T
E
D**

Patent No 6025190

```

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1098..1283
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1620..1733
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2047..2174
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2267..2425
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2567..2751
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2828..2921
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3031..3765
; FEATURE:
; NAME/KEY: -
; LOCATION: 1433..1434
; OTHER INFORMATION: /note= "interruption of sequence data"
; FEATURE:
; NAME/KEY: -
; LOCATION: 2208..2209
; OTHER INFORMATION: /note= "interruption of sequence data"
; FEATURE:
; NAME/KEY: -
; LOCATION: 2781..2782
; OTHER INFORMATION: /note= "interruption of sequence data"
; FEATURE:
; NAME/KEY: -
; LOCATION: 2956..2957
; OTHER INFORMATION: /note= "interruption of sequence data"
;
US-08-410-540-3

Query Match 44.7%; Score 67; DB 2; Length 4016;
Best Local Similarity 100.0%; Pred. No. 8e-15;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 51 GTGATCCCGCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
      |||||||
Db 3939 GTGATCCCGCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 3998
QY 111 GGGGGCC 117
      |||||||
Db 3999 GGGGGCC 4005

RESULT 6
US-08-992-334-2/c
; Sequence 2, Application US/08992334
; Patent No. 5919678
; GENERAL INFORMATION:
; APPLICANT: Gruss, Alexandra
; APPLICANT: Maquin, Emmanuelle
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE
; TITLE OF INVENTION: PLASMID
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christie Parker & Hale, LLP
; STREET: 350 West Colorado Boulevard, Suite 500
; CITY: Pasadena
; STATE: California
; COUNTRY: United States

```

```
;
;
; ZIP: 91105
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992.334
; FILING DATE: 17-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,752
; FILING DATE: 24-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00248
; FILING DATE: 12-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR FR92/03034
; FILING DATE: 13-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Prout, D. Bruce
; REGISTRATION NUMBER: 20958
; REFERENCE/DOCKET NUMBER: C93:31779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626) 795-9900
; TELEFAX: (626) 577-8800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5234 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-992-334-2

Query Match 44.7%; Score 67; DB 4; Length 5234;
Best Local Similarity 100.0%; Pred. No. 8.8e-15;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
Db 4983 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4924

QY 111 GGGGGCC 117
Db 4923 GGGGGCC 4917

RESULT 7
US-08-302-752-2/c
; Sequence 2, Application US/08302752
; Patent No. 6025190
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: THERMOSENSIBLE PLASMID
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302.752
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9203034
; FILING DATE: 13-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR/93/00248
; FILING DATE: 12-MAR-1993
; INFORMATION FOR SEQ ID NO: 2:
```

```
;
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5234 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-302-752-2

Query Match 44.7%; Score 67; DB 5; Length 5234;
Best Local Similarity 100.0%; Pred. No. 8.8e-15;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
Db 4983 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 4924

QY 111 GGGGGCC 117
Db 4923 GGGGGCC 4917

RESULT 8
US-08-992-334-3/c
; Sequence 3, Application US/08992334
; Patent No. 5919678
; GENERAL INFORMATION:
; APPLICANT: Gruss, Alexandra
; APPLICANT: Maguin, Emmanuelle
; TITLE OF INVENTION: METHODS FOR USING A TEMPERATURE-SENSITIVE
; MOLECULE TYPE: PLASMID
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christie Parker & Hale, LLP
; STREET: 350 West Colorado Boulevard, Suite 500
; CITY: Pasadena
; STATE: California
; COUNTRY: United States
; ZIP: 91105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992.334
; FILING DATE: 17-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,752
; FILING DATE: 24-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00248
; FILING DATE: 12-MAR-1993
; PRIOR APPLICATION DATA: FR FR92/03034
; FILING DATE: 13-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Prout, D. Bruce
; REGISTRATION NUMBER: 20958
; REFERENCE/DOCKET NUMBER: C93:31779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626) 795-9900
; TELEFAX: (626) 577-8800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-992-334-3
```

```
Query Match          44.7%; Score 67; DB 4; Length 6722;
Best Local Similarity 100.0%; Pred. No. 9.6e-15;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
DB 6471 GTGGATCCCCGGGTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 6412

QY 111 GGGGGCC 117
DB 6411 GGGGGCC 6405

RESULT 9
US-08-302-752-3/c
; Sequence 3, Application US/08302752
; Patent No. 6025190
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: THERMOSENSIBLE PLASMID
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,752
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9203034
; FILING DATE: 13-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR/93/00248
; FILING DATE: 12-MAR-1993
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-302-752-3

Query Match          44.7%; Score 67; DB 5; Length 6722;
Best Local Similarity 100.0%; Pred. No. 9.6e-15;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
DB 6471 GTGGATCCCCGGGTGCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 6412

QY 111 GGGGGCC 117
DB 6411 GGGGGCC 6405

RESULT 10
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixaldo, Marmelstein, Murray & Oram LLP
```

```
STREET: 655 Fifteenth Street N.W. Suite 330
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,610
FILING DATE: 07-MAR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 44 31 973.8
FILING DATE: 08-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 195 03 952.1
FILING DATE: 07-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/03175
FILING DATE: 10-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Berman, Richard J.
REGISTRATION NUMBER: 39,105
REFERENCE/DOCKET NUMBER: P1614-7007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)638-5000
TELEFAX: (202)638-4810
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 9318 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA
US-08-793-610-6

Query Match          44.1%; Score 66.2; DB 3; Length 9318;
Best Local Similarity 95.8%; Pred. No. 2.1e-14;
Matches 68; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 47 TGACGTGATCCCCCGGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 106
DB 5650 TAATGGGATCCCCCGGGCTCGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTC 5709

QY 107 GAGGGGGGCC 117
DB 5710 GAGGGGGGCC 5720

RESULT 11
US-08-463-115-56
; Sequence 56, Application US/08463115
; Patent No. 5703221
; GENERAL INFORMATION:
; APPLICANT: WILLIAM JOHN MARTIN
; TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
; TITLE OF INVENTION: AND RELATED VACCINES
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
```

```

; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,115
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/157,811
; FILING DATE: No. 5703221ember 23, 1993
; APPLICATION NUMBER: 07/887,502
; FILING DATE: May 22, 1992
; APPLICATION NUMBER: 07/704,814
; FILING DATE: May 23, 1991
; APPLICATION NUMBER: 07/763,039
; FILING DATE: September 20, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 213/301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 685 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION:
; US-08-463-115-56

Query Match 41.6%; Score 62.4; DB 2; Length 685;
Best Local Similarity 94.0%; Pred. No. 1.9e-13;
Matches 63; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
DB 150 GTGGATCCCCGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 209

QY 111 GGGGGCC 117
DB 210 GGGGNC 216

RESULT 12
US-08-463-388-56
; Sequence 56, Application US/08465388
; Patent No. 5753488
; GENERAL INFORMATION:
; APPLICANT: WILLIAM JOHN MARTIN
; TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
; TITLE OF INVENTION: AND RELATED VACCINES
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/465,388
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/157,811
; FILING DATE: No. 5753488ember 23, 1993
; APPLICATION NUMBER: 07/887,502
; FILING DATE: May 22, 1992
; APPLICATION NUMBER: 07/704,814
; FILING DATE: May 23, 1991
; APPLICATION NUMBER: 07/763,039
; FILING DATE: September 20, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 213/300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 685 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION:
; US-08-465-388-56

Query Match 41.6%; Score 62.4; DB 2; Length 685;
Best Local Similarity 94.0%; Pred. No. 1.9e-13;
Matches 63; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
DB 150 GTGGATCCCCGGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 209

QY 111 GGGGGCC 117
DB 210 GGGGNC 216

RESULT 13
US-08-204-675-1
; Sequence 1, Application US/08204675
; Patent No. 5677170
; GENERAL INFORMATION:
; APPLICANT: Devine, Scott E.
; APPLICANT: Boeke, Jef D.
; APPLICANT: Braiterman, Lelita T.
; TITLE OF INVENTION: In Vitro Transposition of Artificial
; TITLE OF INVENTION: Transposons
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie, and Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,675
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 435

```


ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.45501
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202.508.9100
TELEFAX: 202.508.9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4164 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: PAT-1
US-08-204-675-1

Query Match 39.3%; Score 59; DB 1; Length 4164;
Best Local Similarity 100.0%; Pred. No. 6e-12;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 109
|||||
DB 2044 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102

RESULT 14
US-08-660-754-1
Sequence 1, Application US/08660754
Patent No. 5843772
GENERAL INFORMATION:
APPLICANT: Devine, Scott E.
APPLICANT: Boeke, Jef D.
APPLICANT: Braiterman, Lelita T.
TITLE OF INVENTION: In Vitro Transposition of Artificial
TITLE OF INVENTION: Transposons
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie, and Beckett
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,754
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,675
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.45501
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202.508.9100
TELEFAX: 202.508.9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4164 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: PAT-1
US-08-660-754-1

Query Match 39.3%; Score 59; DB 3; Length 4164;
Best Local Similarity 100.0%; Pred. No. 6e-12;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 109
|||||
DB 2044 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG 2102

RESULT 15
US-08-796-364-1
Sequence 1, Application US/08796364
Patent No. 5968785
GENERAL INFORMATION:
APPLICANT: Devine, Scott E.
APPLICANT: Boeke, Jef D.
APPLICANT: Braiterman, Lelita T.
TITLE OF INVENTION: In Vitro Transposition of Artificial
TITLE OF INVENTION: Transposons
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie, and Beckett
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/796,364
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,675
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.45501
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202.508.9100
TELEFAX: 202.508.9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4164 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: PAT-1
US-08-796-364-1

Query Match 39.3%; Score 59; DB 4; Length 4164;
Best Local Similarity 100.0%; Pred. No. 6e-12;

```
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	51	GTGGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG	109
Db	2044	GTGGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG	2102

Search completed: May 29, 2000, 22:09:16
Job time: 38757 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2000, 09:48:32 ; Search time 2276.24 Seconds
(without alignments)
294.022 Million cell updates/sec

Title: US-08-935-377-9

Perfect score: 150

Sequence: 1 GGCAAAATGAAAACTA.....TTGTTTTGTGGCCCGGCC 150

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5142629 seqs, 2230885800 residues

Total number of hits satisfying chosen parameters: 10285240

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA.*

1: /cgnl_6/ptodata/1/pna/PCTUS_COMB.seq.*
2: /cgnl_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgnl_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgnl_6/ptodata/1/pna/US08_COMB.seq.*
5: /cgnl_6/ptodata/1/pna/US081A_COMB.seq.*
6: /cgnl_6/ptodata/1/pna/US081B_COMB.seq.*
7: /cgnl_6/ptodata/1/pna/US081C_COMB.seq.*
8: /cgnl_6/ptodata/1/pna/US081D_COMB.seq.*
9: /cgnl_6/ptodata/1/pna/US082A_COMB.seq.*
10: /cgnl_6/ptodata/1/pna/US082B_COMB.seq.*
11: /cgnl_6/ptodata/1/pna/US082C_COMB.seq.*
12: /cgnl_6/ptodata/1/pna/US083A_COMB.seq.*
13: /cgnl_6/ptodata/1/pna/US083B_COMB.seq.*
14: /cgnl_6/ptodata/1/pna/US084A_COMB.seq.*
15: /cgnl_6/ptodata/1/pna/US084B_COMB.seq.*
16: /cgnl_6/ptodata/1/pna/US084C_COMB.seq.*
17: /cgnl_6/ptodata/1/pna/US085_COMB.seq.*
18: /cgnl_6/ptodata/1/pna/US086_COMB.seq.*
19: /cgnl_6/ptodata/1/pna/US087A_COMB.seq.*
20: /cgnl_6/ptodata/1/pna/US087B_COMB.seq.*
21: /cgnl_6/ptodata/1/pna/US087C_COMB.seq.*
22: /cgnl_6/ptodata/1/pna/US088A_COMB.seq.*
23: /cgnl_6/ptodata/1/pna/US088B_COMB.seq.*
24: /cgnl_6/ptodata/1/pna/US088C_COMB.seq.*
25: /cgnl_6/ptodata/1/pna/US088D_COMB.seq.*
26: /cgnl_6/ptodata/1/pna/US089A_COMB.seq.*
27: /cgnl_6/ptodata/1/pna/US089B_COMB.seq.*
28: /cgnl_6/ptodata/1/pna/US089C_COMB.seq.*
29: /cgnl_6/ptodata/1/pna/US089D_COMB.seq.*
30: /cgnl_6/ptodata/1/pna/US089E_COMB.seq.*
31: /cgnl_6/ptodata/1/pna/US089F_COMB.seq.*
32: /cgnl_6/ptodata/1/pna/US089G_COMB.seq.*
33: /cgnl_6/ptodata/1/pna/US089H_COMB.seq.*
34: /cgnl_6/ptodata/1/pna/US089I_COMB.seq.*
35: /cgnl_6/ptodata/1/pna/US090A_COMB.seq.*
36: /cgnl_6/ptodata/1/pna/US090B_COMB.seq.*
37: /cgnl_6/ptodata/1/pna/US090C_COMB.seq.*
38: /cgnl_6/ptodata/1/pna/US090D_COMB.seq.*
39: /cgnl_6/ptodata/1/pna/US090E_COMB.seq.*
40: /cgnl_6/ptodata/1/pna/US091A_COMB.seq.*
41: /cgnl_6/ptodata/1/pna/US091B_COMB.seq.*
42: /cgnl_6/ptodata/1/pna/US091C_COMB.seq.*
43: /cgnl_6/ptodata/1/pna/US092A_COMB.seq.*

44: /cgnl_6/ptodata/1/pna/US092B_COMB.seq.*
45: /cgnl_6/ptodata/1/pna/US092C_COMB.seq.*
46: /cgnl_6/ptodata/1/pna/US092D_COMB.seq.*
47: /cgnl_6/ptodata/1/pna/US092E_COMB.seq.*
48: /cgnl_6/ptodata/1/pna/US093A_COMB.seq.*
49: /cgnl_6/ptodata/1/pna/US093B_COMB.seq.*
50: /cgnl_6/ptodata/1/pna/US093C_COMB.seq.*
51: /cgnl_6/ptodata/1/pna/US093D_COMB.seq.*
52: /cgnl_6/ptodata/1/pna/US093E_COMB.seq.*
53: /cgnl_6/ptodata/1/pna/US093F_COMB.seq.*
54: /cgnl_6/ptodata/1/pna/US094A_COMB.seq.*
55: /cgnl_6/ptodata/1/pna/US094B_COMB.seq.*
56: /cgnl_6/ptodata/1/pna/US094C_COMB.seq.*
57: /cgnl_6/ptodata/1/pna/US095_COMB.seq.*
58: /cgnl_6/ptodata/1/pna/US6000_COMB.seq.*
59: /cgnl_6/ptodata/1/pna/US6001A_COMB.seq.*
60: /cgnl_6/ptodata/1/pna/US6001B_COMB.seq.*
61: /cgnl_6/ptodata/1/pna/US6002A_COMB.seq.*
62: /cgnl_6/ptodata/1/pna/US6002B_COMB.seq.*
63: /cgnl_6/ptodata/1/pna/US6002C_COMB.seq.*
64: /cgnl_6/ptodata/1/pna/US6003A_COMB.seq.*
65: /cgnl_6/ptodata/1/pna/US6003B_COMB.seq.*
66: /cgnl_6/ptodata/1/pna/US6004A_COMB.seq.*
67: /cgnl_6/ptodata/1/pna/US6004B_COMB.seq.*
68: /cgnl_6/ptodata/1/pna/US6004C_COMB.seq.*
69: /cgnl_6/ptodata/1/pna/US6005_COMB.seq.*
70: /cgnl_6/ptodata/1/pna/US6006_COMB.seq.*
71: /cgnl_6/ptodata/1/pna/US6007_COMB.seq.*
72: /cgnl_6/ptodata/1/pna/US6008A_COMB.seq.*
73: /cgnl_6/ptodata/1/pna/US6008B_COMB.seq.*
74: /cgnl_6/ptodata/1/pna/US6008C_COMB.seq.*
75: /cgnl_6/ptodata/1/pna/US6009A_COMB.seq.*
76: /cgnl_6/ptodata/1/pna/US6009B_COMB.seq.*
77: /cgnl_6/ptodata/1/pna/US6010A_COMB.seq.*
78: /cgnl_6/ptodata/1/pna/US6010B_COMB.seq.*
79: /cgnl_6/ptodata/1/pna/US6011A_COMB.seq.*
80: /cgnl_6/ptodata/1/pna/US6011B_COMB.seq.*
81: /cgnl_6/ptodata/1/pna/US6012A_COMB.seq.*
82: /cgnl_6/ptodata/1/pna/US6012B_COMB.seq.*
83: /cgnl_6/ptodata/1/pna/US6013A_COMB.seq.*
84: /cgnl_6/ptodata/1/pna/US6013B_COMB.seq.*
85: /cgnl_6/ptodata/1/pna/US6014A_COMB.seq.*
86: /cgnl_6/ptodata/1/pna/US6014B_COMB.seq.*
87: /cgnl_6/ptodata/1/pna/US6014C_COMB.seq.*
88: /cgnl_6/ptodata/1/pna/US6015_COMB.seq.*
89: /cgnl_6/ptodata/1/pna/US6016A_COMB.seq.*
90: /cgnl_6/ptodata/1/pna/US6016B_COMB.seq.*
91: /cgnl_6/ptodata/1/pna/US6016C_COMB.seq.*
92: /cgnl_6/ptodata/1/pna/US6017A_COMB.seq.*
93: /cgnl_6/ptodata/1/pna/US6017B_COMB.seq.*
94: /cgnl_6/ptodata/1/pna/US6017C_COMB.seq.*
95: /cgnl_6/ptodata/1/pna/US6018A_COMB.seq.*
96: /cgnl_6/ptodata/1/pna/US6018B_COMB.seq.*
97: /cgnl_6/ptodata/1/pna/US6018C_COMB.seq.*
98: /cgnl_6/ptodata/1/pna/US6019_COMB.seq.*
99: /cgnl_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
100: /cgnl_6/ptodata/1/pna/US06_NEW_COMB.seq.*
101: /cgnl_6/ptodata/1/pna/US07_NEW_COMB.seq.*
102: /cgnl_6/ptodata/1/pna/US08_NEW_COMB.seq.*
103: /cgnl_6/ptodata/1/pna/US09_NEW_COMB.seq.*
104: /cgnl_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	150.	100.0	150	28	US-08-935-377-9	Sequence 9, Appli
2	138	92.0	149	28	US-08-935-377-8	Sequence 8, Appli


```
Query Match      92.0%; Score 138; DB 28; Length 149;
Best Local Similarity 99.3%; Pred. No. 1.3e-37;
Matches 149; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 GGCCAAAATTGAAAACCTAGATCTATTATTGACGGCGCGCCCATGACGTGGATCCCC 60
    |||||||
Db 1 GGCCAAAATTGAAAACCTAGATCTATTATTGACGGCGCGCCCATGAGTGGATCCCC 59
    |||||||
Qy 61 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCCTAA 120
    |||||||
Db 60 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCCTAA 119
    |||||||
Qy 121 CTAACATAATTTGTTTTGTGGCCCGCC 150
    |||||||
Db 120 CTAACATAATTTGTTTTGTGGCCCGCC 149
    |||||||

RESULT 3
US-08-935-377-7
; Sequence 7, Application US/08935377
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; TITLE OF INVENTION: T Cells Specific for Target Antigens and
; TITLE OF INVENTION: Vaccines Based Thereon
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 148 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-935-377-7

Query Match      90.7%; Score 136; DB 28; Length 148;
Best Local Similarity 98.7%; Pred. No. 6.4e-37;
Matches 148; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

Qy 1 GGCCAAAATTGAAAACCTAGATCTATTATTGACGGCGCGCCCATGACGTGGATCCCC 60
    |||||||
Db 1 GGCCAAAATTGAAAACCTAGATCTATTATTGACGGCGCGCCCATG--GTGGATCCCC 58
    |||||||
Qy 61 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCCTAA 120
    |||||||
Db 59 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCCTAA 118
    |||||||
Qy 121 CTAACATAATTTGTTTTGTGGCCCGCC 150
    |||||||
```

```
Db 119 CTAACATAATTTGTTTTGTGGCCCGCC 148
    |||||||

RESULT 4
US-08-935-377-6
; Sequence 6, Application US/08935377
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; TITLE OF INVENTION: T Cells Specific for Target Antigens and
; TITLE OF INVENTION: Vaccines Based Thereon
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1821.0010000/EKS/CMB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-935-377-6

Query Match      86.7%; Score 130; DB 28; Length 145;
Best Local Similarity 96.7%; Pred. No. 7.5e-35;
Matches 145; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

Qy 1 GGCCAAAATTGAAAACCTAGATCTATTATTGACGGCGCGCCCATGACGTGGATCCCC 60
    |||||||
Db 1 GGCCAAAATTGAAAACCTAGATCTATTATTGACGGCGCGCCCATGACGTGGATCCCC 55
    |||||||
Qy 61 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCCTAA 120
    |||||||
Db 56 CGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGCCTAA 115
    |||||||
Qy 121 CTAACATAATTTGTTTTGTGGCCCGCC 150
    |||||||
Db 116 CTAACATAATTTGTTTTGTGGCCCGCC 145
    |||||||

RESULT 5
US-08-431-048A-121/c
; Sequence 121, Application US/08431048A
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Administrator, TESTA, HURWITZ &
```

```

; ; ADDRESSSEE: THIBEAULT, LLP
; ; STREET: 125 High Street
; ; CITY: BOSTON
; ; STATE: MA
; ; COUNTRY: USA
; ; ZIP: 02110
; ; COMPUTER READABLE FORM:
; ; MEDIUM TYPE: Floppy disk
; ; COMPUTER: IBM PC compatible
; ; OPERATING SYSTEM: PC-DOS/MS-DOS
; ; SOFTWARE: Patent In Release #1.0, Version #1.30
; ; CURRENT APPLICATION DATA:
; ; APPLICATION NUMBER: US/08/431.048A
; ; FILING DATE: 28-APR-1995
; ; CLASSIFICATION: 800
; ; ATTORNEY/AGENT INFORMATION:
; ; NAME: Twomey, Michael J
; ; REGISTRATION NUMBER: 38,349
; ; REFERENCE/DOCKET NUMBER: CAN-002
; ; TELECOMMUNICATION INFORMATION:
; ; TELEPHONE: (617) 248-7100
; ; TELEFAX: (617) 248-7000
; ; INFORMATION FOR SEQ ID NO: 121:
; ; LENGTH: 793 base pairs
; ; TYPE: nucleic acid
; ; STRANDEDNESS: single
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: DNA (genomic)
; ; US-08-431-048A-121
; ;
; ; Query Match 46.5%; Score 69.8; DB 15; Length 793;
; ; Best Local Similarity 82.5%; Pred. No. 7.9e-14;
; ; Matches 80; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
; ;
; ; YQ 39 GGCGCCATCAGTGGATCCCCGCGGTGCAGGAATTCGATCAAGCTTATCGATACCG 98
; ; Db 130 GGAGCGCGGGCGGATCCCCGCGGTGCAGGAATTCGATCAAGCTTATCGATACCG 71
; ;
; ; YQ 99 TCGACCTCGAGGGGGGCCCTACTAACTTTTGT 135
; ; Db 70 TCGACCTCGAGGGGGGCCCTACTAACTTTTGT 34
; ;
; ; RESULT 6
; ; US-08-496-841-121/c
; ; Sequence 121, Application US/08496841
; ; GENERAL INFORMATION:
; ; APPLICANT: St. George-Hyslop, Peter
; ; APPLICANT: Rommens, Johanna M.
; ; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; ; TO ALZHEIMER'S DISEASE
; ; NUMBER OF SEQUENCES: 160
; ; CORRESPONDENCE ADDRESS:
; ; ADDRESSEE: Testa, Hurwitz & Thibault
; ; STREET: High Street Tower - 125 High Street
; ; CITY: Boston
; ; STATE: Massachusetts
; ; COUNTRY: U.S.A.
; ; ZIP: 02110
; ; COMPUTER READABLE FORM:
; ; MEDIUM TYPE: Floppy disk
; ; COMPUTER: IBM PC compatible
; ; OPERATING SYSTEM: PC-DOS/MS-DOS
; ; SOFTWARE: Patent In Release #1.0, Version #1.30
; ; CURRENT APPLICATION DATA:
; ; APPLICATION NUMBER: US/08/496.841
; ; FILING DATE: 28-JUN-1995
; ; CLASSIFICATION: 424
; ; ATTORNEY/AGENT INFORMATION:
; ; NAME: Twomey, Michael J
; ; REGISTRATION NUMBER: 38,349
; ; REFERENCE/DOCKET NUMBER: CAN-002
; ; TELECOMMUNICATION INFORMATION:
; ; TELEPHONE: (617) 248-7100
; ; TELEFAX: (617) 248-7000
; ; INFORMATION FOR SEQ ID NO: 121:
; ; LENGTH: 793 base pairs
; ; TYPE: nucleic acid
; ; STRANDEDNESS: single
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: DNA (genomic)
; ; US-08-496-841-121
; ;
; ; Query Match 46.5%; Score 69.8; DB 15; Length 793;
; ; Best Local Similarity 82.5%; Pred. No. 7.9e-14;
; ; Matches 80; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
; ;
; ; YQ 39 GGCGCCATCAGTGGATCCCCGCGGTGCAGGAATTCGATCAAGCTTATCGATACCG 98
; ; Db 130 GGAGCGCGGGCGGATCCCCGCGGTGCAGGAATTCGATCAAGCTTATCGATACCG 71
; ;
; ; YQ 99 TCGACCTCGAGGGGGGCCCTACTAACTTTTGT 135
; ; Db 70 TCGACCTCGAGGGGGGCCCTACTAACTTTTGT 34
; ;
; ; RESULT 7
; ; US-08-509-359-121/c
; ; Sequence 121, Application US/08509359
; ; GENERAL INFORMATION:
; ; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; ; APPLICANT: ROMMENS, JOHANNA M
; ; APPLICANT: FRASER, PAUL E
; ; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; ; TO ALZHEIMER'S DISEASE
; ; NUMBER OF SEQUENCES: 160
; ; CORRESPONDENCE ADDRESS:
; ; ADDRESSEE: Testa, Hurwitz & Thibault
; ; STREET: High Street Tower - 125 High Street
; ; CITY: Boston
; ; STATE: Massachusetts
; ; COUNTRY: U.S.A.
; ; ZIP: 02110
; ; COMPUTER READABLE FORM:
; ; MEDIUM TYPE: Floppy disk
; ; COMPUTER: IBM PC compatible
; ; OPERATING SYSTEM: PC-DOS/MS-DOS
; ; SOFTWARE: Patent In Release #1.0, Version #1.30
; ; CURRENT APPLICATION DATA:
; ; APPLICATION NUMBER: US/08/509,359
; ; FILING DATE: 31-JUL-1995
; ; CLASSIFICATION: 800
; ; ATTORNEY/AGENT INFORMATION:
; ; NAME: Twomey, Michael J
; ; REGISTRATION NUMBER: 38349
; ; TELECOMMUNICATION INFORMATION:
; ; TELEPHONE: (617) 248-7362
; ; TELEFAX: (617) 248-7100
; ; INFORMATION FOR SEQ ID NO: 121:
; ; LENGTH: 793 base pairs
; ; TYPE: nucleic acid
; ; STRANDEDNESS: single
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: DNA (genomic)
; ; US-08-509-359-121
; ;
; ; Query Match 46.5%; Score 69.8; DB 17; Length 793;
; ; Best Local Similarity 82.5%; Pred. No. 7.9e-14;
; ; Matches 80; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```


CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 793 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-124-698-121

Query Match 46.5%; Score 69.8; DB 40; Length 793;
Best Local Similarity 82.5%; Pred. No. 7.9e-14;
Matches 80; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 39 GGCCGCGGATGAGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 98
DB 130 GGAGCCGCGGGCGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 71
QY 99 TCGACCTCGAGGGGGGCGCTAACTAACTAATTTGTT 135
DB 70 TCGACCTCGAGGGGGGCGCTAACTAACTAATTTGTT 34

RESULT 11
US-09-127-480-121/c
SEQUENCE 121, Application US/09127480
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/127.480
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 793 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-09-127-480-121

Query Match 46.5%; Score 69.8; DB 40; Length 793;
Best Local Similarity 82.5%; Pred. No. 7.9e-14;
Matches 80; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 39 GGCCGCGGATGAGTGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 98
DB 130 GGAGCCGCGGGCGGATCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCG 71
QY 99 TCGACCTCGAGGGGGGCGCTAACTAACTAATTTGTT 135
DB 70 TCGACCTCGAGGGGGGCGCTAACTAACTAATTTGTT 34

RESULT 12
US-08-377-383-4
SEQUENCE 4, Application US/08377383
GENERAL INFORMATION:
APPLICANT: HOSOI, Shigeru, FUKAMI, Tadashi, HIYOSHI, Makiko
TITLE OF INVENTION: Method of Determining Base Sequence of
TITLE OF INVENTION: Nucleic Acid
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cushman Darby & Cushman
STREET: Ninth Floor, 1100 New York Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 720Kb storage
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1 or ASCII editors
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/377,383
FILING DATE: 25-JAN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP6971/1994
FILING DATE: 26-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Glenn J. Perry
REGISTRATION NUMBER: 28458
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)861-3000
TELEFAX: (202)822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid Synthetic DNA
US-08-377-383-4

Query Match 46.0%; Score 69; DB 13; Length 224;
Best Local Similarity 88.2%; Pred. No. 1e-13;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTTCGATACCG 110
DB 88 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTTCGATACCG 147
QY 111 GGGGGCGCTAACTAACTAATTTGTT 135
DB 148 GGGGGCGGCTACCCAGCTTTGTT 172
RESULT 13

US-08-693-573-4
; Sequence 4, Application US/08693573
; GENERAL INFORMATION:
; APPLICANT: HOSOI, Shigeru,
; APPLICANT: FUKAMI, Tadashi,
; APPLICANT: KOJIMA, Makiko
; TITLE OF INVENTION: Method of Determining Base Sequence of
; Nucleic Acid
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY CUSHMAN
; ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: Ninth Floor, 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 720Kb storage
; COMPUTER: IBM PC/AT/AT compatibles
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1 or ASCII editors
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,573
; FILING DATE: 07-AUG-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/377,383
; FILING DATE: 25-JAN-1995
; APPLICATION NUMBER: JP6971/1994
; FILING DATE: 26-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Thomas G. Wiseman
; REGISTRATION NUMBER: 35046
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)861-3000
; TELEFAX: (202)822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other nucleic acid Synthetic DNA
; US-08-693-573-4

Query Match 46.0%; Score 69; DB 18; Length 224;
Best Local Similarity 88.2%; Pred. No. 1e-13;
Matches 75; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 51 GTGATCCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 88 GTGATCCCCGGGCTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 147
QY 111 GGGGGCCCTAACTAACTAATTGTT 135
|||||
Db 148 GGGGGCCCGGTACCGACCTTTGTT 172
RESULT 14
US-061-998-2/c
; Sequence 2, Application US/60061998
; GENERAL INFORMATION:
; APPLICANT: LAGACE, ROBERT E.
; APPLICANT: CORLEY, NEIL C.
; APPLICANT: RUSSO, FRANK D.
; APPLICANT: HANN, AMY L.
; APPLICANT: HEATH, JOE D.
; APPLICANT: FINNEY, GREGORY L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE
; GENOME, FRAGMENTS THEREOF, AND USES THEREOF

NUMBER OF SEQUENCES: 797
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/061,998
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PM-0006-2P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
CLONE: SPN1C002
US-60-061-998-2
Query Match 45.2%; Score 67.8; DB 70; Length 1172;
Best Local Similarity 75.7%; Pred. No. 4.4e-13;
Matches 84; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 5 AAAAATTGAAAACTAGATCTATTATTGCACGCGCGCCGCGCATGACGTGATCCCCGGG 64
|||||
Db 111 AAGAAAAATAATGGAATATCTGAATTCGCGTACCAAGTAAGGCGGATCCCCGGG 52
QY 65 CTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGG 115
|||||
Db 51 CTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGG 1
RESULT 15
US-09-072-433-31/c
; Sequence 31, Application US/09072433
; GENERAL INFORMATION:
; APPLICANT: Coschigano, Peter W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITRATION OF BIOMEDIATION
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,433
; FILING DATE: 04-MAY-1998
; CLASSIFICATION: 514

```
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OHU-03344
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 147 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..147
US-09-072-433-31
```

```
Query Match 44.7%; Score 67; DB 38; Length 147;
Best Local Similarity 100.0%; Pred. No. 4.4e-13;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 47

Qy 111 GGGGGCC 117
   |||||
Db 46 GGGGGCC 40
```

```
Search completed: May 30, 2000, 09:48:32
Job time: 50182 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: May 29, 2000, 21:13:16 ; Search time 2192.43 Seconds
(without alignments)
277.310 Million cell updates/sec

Title: us-08-935-377-9
Perfect score: 150
Sequence: 1 GCCCAAAATGAAAAACTA.....TTGTTTTGTGGCCCGGCC 150

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues
Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *

44: gb_est25: *
45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: em_est20: *
50: em_est21: *
51: em_est22: *
52: gb_est30: *
53: gb_est31: *
54: gb_est32: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: gb_gss1: *
83: gb_gss2: *
84: gb_gss3: *
85: gb_gss4: *
86: em_gss1: *
87: em_gss2: *
88: em_gss3: *
89: em_gss4: *
90: gb_gss5: *
91: gb_gss6: *
92: em_gss5: *
93: em_gss6: *
94: gb_gss7: *
95: gb_gss8: *
96: gb_gss9: *
97: em_gss7: *
98: em_gss8: *
99: gb_gss11: *
100: gb_gss10: *
101: em_gss9: *
102: em_gss10: *
103: em_gss11: *
104: em_gss12: *
105: gb_gss12: *
106: gb_gss13: *
107: gb_gss14: *
108: gb_gss15: *
109: gb_gss16: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	70	46.7	682	96	AQ074693	AQ074693 CIT-HSP-2
C 2	68	45.3	608	96	AQ009167	AQ009167 CIT-HFP-2
C 3	67.4	44.9	528	84	B69688	B69688 CIT978SK-A-
C 4	67	44.7	108	84	B54292	B54292 CIT-HSP-201
C 5	67	44.7	141	44	AQ037147	AQ037147 AU037147
C 6	67	44.7	141	96	B95279	B95279 CIT-HSP-217
C 7	67	44.7	144	84	B80294	B80294 CIT-HSP-204
C 8	67	44.7	169	96	AQ013255	AQ013255 CIT-HSP-2
C 9	67	44.7	175	96	AQ041198	AQ041198 CIT-HSP-2
C 10	67	44.7	259	20	T48593	T48593 ph0f4_19/IT
C 11	67	44.7	275	27	AA052885	AA052885 T351_110
C 12	67	44.7	386	84	B74162	B74162 CIT-HSP-202
C 13	67	44.7	462	96	AQ014512	AQ014512 CIT-HSP-2
C 14	67	44.7	519	84	B65534	B65534 CIT-HSP-202
C 15	67	44.7	547	96	AQ012629	AQ012629 CIT-HSP-2
C 16	67	44.7	552	96	AQ012216	AQ012216 CIT-HSP-2
C 17	67	44.7	575	96	AQ012221	AQ012221 CIT-HSP-2
C 18	67	44.7	596	96	AQ009165	AQ009165 CIT-HFP-2
C 19	67	44.7	607	84	B64363	B64363 CIT-HSP-202
C 20	67	44.7	632	84	B70299	B70299 CIT-HSP-202
C 21	67	44.7	638	96	AQ016737	AQ016737 CIT-HSP-2
C 22	67	44.7	647	84	B68775	B68775 CIT-HSP-205
C 23	67	44.7	666	84	B80296	B80296 CIT-HSP-204
C 24	67	44.7	700	96	AQ009153	AQ009153 CIT-HSP-2
C 25	66.2	44.1	576	38	AA752172	AA752172 96B80243
C 26	66	44.0	541	96	AQ074695	AQ074695 CIT-HSP-2
C 27	65.4	43.6	513	96	AQ008437	AQ008437 CIT-HSP-2
C 28	65.4	43.6	541	96	AQ014519	AQ014519 CIT-HSP-2
C 29	62.8	41.9	134	84	B72088	B72088 CIT978SK-13
C 30	62.2	41.5	233	96	AQ012365	AQ012365 CIT-HSP-2
C 31	62	41.3	294	20	T26395	T26395 AB161A4F in
C 32	60.8	40.5	133	84	B26269	B26269 F10B5TF IGF
C 33	60	40.0	100	43	A1239372	A1239372 SMOVAFCAP
C 34	57.8	38.5	145	84	B53695	B53695 CIT-HSP-201
C 35	57.8	38.5	630	96	AQ003238	AQ003238 RFL111-19
C 36	57.4	38.3	603	38	AA752115	AA752115 96BS0066
C 37	57.2	38.1	298	82	RP068746	RP068746 R1ftia pach
C 38	56.8	37.9	97	41	AI058116	AI058116 SWMACG15
C 39	56	37.3	144	84	B65391	B65391 CIT-HSP-202
C 40	56	37.3	471	96	AQ024018	AQ024018 CpgR0359A
C 41	56	37.3	499	96	AQ023564	AQ023564 CpgR0482A
C 42	56	37.3	547	96	AQ023548	AQ023548 CpgR0051A
C 43	56	37.3	626	96	AQ024077	AQ024077 CpgR0415A
C 44	55.4	36.9	502	96	AQ023674	AQ023674 CpgR0007B
C 45	55	36.7	149	96	AQ044029	AQ044029 CIT-HSP-2

ALIGNMENTS

RESULT 1	AQ074693	682 bp	DNA	GSS	20-AUG-1998
LOCUS	CIT-HSP-2301L23.TF CIT-HSP Homo sapiens genomic clone 2301L23, genomic survey sequence.				
DEFINITION	genomic survey sequence.				
ACCESSION	AQ074693				
VERSION	AQ074693.1 GI:3436811				
KEYWORDS	GSS.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 682) Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.				
TITLE	Use of a random human BAC End Sequence Database for Sequence-Ready Map Building				
JOURNAL	Unpublished (1998)				
COMMENT	Contact: Mark Adams				

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES

Location/Qualifiers
1..682

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2301L23"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 156 a 187 c 187 g 151 t 1 others
ORIGIN

Query Match 46.7%; Score 70; DB 96; Length 682;
Best Local Similarity 93.6%; Pred. No. 1.4e-13;

Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 51 GTGATCCCCGGGTCAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 78 GTGATCCCCGGGTCAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 19
|||||

QY 111 GGGGGCCTAACTAACTAA 128
|||||

Db 18 GGGGGCCCGAGTACCCAA 1
|||||

RESULT 2
AQ009167/c

LOCUS AQ009167 608 bp DNA GSS 27-JUN-1998
DEFINITION CIT-HFP-2281N13.TF CIT-HSP Homo sapiens genomic clone 2281N13, genomic survey sequence.

ACCESSION AQ009167

VERSION AQ009167.1 GI:3128599

KEYWORDS GSS.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS 1 (bases 1 to 608)
Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J., Shizuya,H., Simon,M. and Venter,J.C.

TITLE Use of a human BAC End Sequence Database for Sequence-Ready Map Building

JOURNAL Unpublished (1997)

COMMENT Other_GSSs: CIT-HSP-2281N13.TR.1 CIT-HSP-2281N13.TF
CIT-HSP-2281N13.TR

Contact: Mark Adams

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html

Seq primer: M13-21;

Class: BAC ends.

Location/Qualifiers

1..608

FEATURES

source

```
/organism="Homo sapiens"
/db_xref="GDB:7147360"
/db_xref="taxon:9606"
/clone="2281N13"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/Note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII"
HindIII"
141 a 164 c 166 g 137 t

BASE COUNT      141 a 164 c 166 g 137 t
ORIGIN

Query Match      45.3%; Score 68; DB 96; Length 608;
Best Local Similarity 100.0%; Pred. No. 6.6e-13;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGG 110
      |||||||
Db 82 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGG 23
      |||||||
QY 111 GGGGGCCT 118
      |||||||
Db 22 GGGGGCCT 15

RESULT 3
B69688/c
LOCUS      B69688      528 bp      DNA      GSS      18-JUN-1998
DEFINITION  CIT978SK-A-448E10.TVC CIT978SK Homo sapiens genomic clone A-448E10,
      genomic survey sequence.
ACCESSION  B69688
VERSION    B69688.1 GI:2708912
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
      Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 528)
AUTHORS    Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
      Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Kim,U.-J.,
      Shizuwa,H., Simon,M. and Venter,J.C.
      Use of a human BAC End Sequence Database for Sequence-Ready Map
      Building
      Unpublished (1997)
JOURNAL
COMMENT     Contact: Mark Adams
      Department of Eukaryotic Genomics
      The Institute for Genomic Research
      9712 Medical Center Dr., Rockville, MD 20850, USA
      Tel: 301 838 0200
      Fax: 301 838 0208
      Email: mdadams@tigr.org
      Clones are available from Research Genetics (info@resgen.com). BAC
      end search page:
      http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
      Seq primer: T7
      Class: BAC ends.

FEATURES             Location/Qualifiers
     source           1..528
     organism="Homo sapiens"
     db_xref="GDB:5299607"
     db_xref="taxon:9606"
     clone="A-448E10"
     clone_lib="CIT978SK"
     sex="Female"
     cell_type="Fibroblast"
     note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
      CalTech Human BAC Library A"

BASE COUNT      118 a 138 c 147 g 125 t
ORIGIN

Query Match      44.9%; Score 67.4; DB 84; Length 528;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGG 110
      |||||||
Db 34 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGG 93
      |||||||
QY 111 GGGGGCCT 117
      |||||||
Db 94 GGGGGCCT 100

RESULT 5
AU037147
```

```
Best Local Similarity 98.6%; Pred. No. 1e-12;
Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGG 110
      |||||||
Db 75 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGG 16
      |||||||
QY 111 GGGGGCCTA 119
      |||||||
Db 15 GGGGGCCCA 7

RESULT 4
B54292
LOCUS      B54292      108 bp      DNA      GSS      20-JUN-1998
DEFINITION  CIT-HSP-2017M1.TR CIT-HSP Homo sapiens genomic clone 2017M1,
      genomic survey sequence.
ACCESSION  B54292
VERSION    B54292.1 GI:2608626
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
      Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 108)
AUTHORS    Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
      Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuwa,H.,
      Simon,M. and Venter,J.C.
      Use of a random BAC End Sequence Database for Sequence-Ready Map
      Building
      Unpublished (1997)
JOURNAL
COMMENT     Contact: Mark Adams
      Department of Eukaryotic Genomics
      The Institute for Genomic Research
      9712 Medical Center Dr., Rockville, MD 20850, USA
      Tel: 301 838 0200
      Fax: 301 838 0208
      Email: mdadams@tigr.org
      Clones are available from Research Genetics (info@resgen.com). BAC
      end search page:
      http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
      Seq primer: M13 Reverse
      Class: BAC ends.

FEATURES             Location/Qualifiers
     source           1..108
     organism="Homo sapiens"
     db_xref="GDB:7044003"
     db_xref="taxon:9606"
     clone="2017M1"
     clone_lib="CIT-HSP"
     sex="Male"
     cell_type="Sperm"
     note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT      20 a 35 c 34 g 19 t
ORIGIN

Query Match      44.7%; Score 67; DB 84; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGG 110
      |||||||
Db 34 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGG 93
      |||||||
QY 111 GGGGGCCT 117
      |||||||
Db 94 GGGGGCCT 100

RESULT 5
AU037147
```

LOCUS AU037147 141 bp mRNA EST 29-MAR-1999
DEFINITION AU037147 Dictyostellium discoideum SS (H.Urushihara) Dictyostellium
discoideum cDNA clone SSB532, mRNA sequence.
ACCESSION AU037147
VERSION AU037147.1 GI:3983900
KEYWORDS EST.
SOURCE Dictyostellium discoideum.
ORGANISM Dictyostellium discoideum.
REFERENCE 1 (bases 1 to 141)
AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mirra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
TITLE The Dictyostellium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development
JOURNAL DNA Res. 5 (6), 335-340 (1998)
MEDLINE 99156227
COMMENT On Jan 19, 1998 this sequence version replaced gi:2153012.
Contact: Hideo Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
Email: d402huesakura.cc.tsukuba.ac.jp
PROJECT = 'Dictyostellium discoideum cDNA project in Japan'.
FEATURES
source
1. .141
/organism="Dictyostellium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SSB532"
/clone_lib="Dictyostellium discoideum SS (H.Urushihara)"
/dev_stage="slug"
BASE COUNT 26 a 40 c 42 g 28 t 5 others
ORIGIN
Query Match 44.7%; Score 67; DB 44; Length 141;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 GTGGATCCCCGGCGTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 69 GTGGATCCCCGGCGTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 128
QY 111 GGGGGCC 117
|||||
Db 129 GGGGGCC 135
RESULT 6
B95279/c 141 bp DNA GSS 25-JUN-1998
LOCUS CIT-HSP-2172N2.TF CIT-HSP Homo sapiens genomic clone 2172N2,
DEFINITION genomic survey sequence.
ACCESSION B95279
VERSION B95279.1 GI:2977616
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 141)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamad@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES
source
1. .144
/organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="taxon:9606"

Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamad@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.
FEATURES
Location/Qualifiers
source
1. .141
/organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="taxon:9606"
/clone="2172N2"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT 24 a 41 c 41 g 35 t
ORIGIN
Query Match 44.7%; Score 67; DB 96; Length 141;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 GTGGATCCCCGGCGTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 73 GTGGATCCCCGGCGTCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 14
QY 111 GGGGGCC 117
|||||
Db 13 GGGGGCC 7
RESULT 7
B80294 144 bp DNA GSS 24-OCT-1998
LOCUS CIT-HSP-2045D19.TR CIT-HSP Homo sapiens genomic clone 2045D19,
DEFINITION genomic survey sequence.
ACCESSION B80294
VERSION B80294.1 GI:2867317
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 144)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Other_GSSs: CIT-HSP-2045D19.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamad@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES
Location/Qualifiers
source
1. .144
/organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="taxon:9606"

```

/clone="2045D19"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/notes="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT      34 a 43 c 41 g 26 t
ORIGIN

Query Match      44.7%; Score 67; DB 84; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
DB 70 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 129
QY 111 GGGGGCC 117
DB 130 GGGGGCC 136

RESULT 8
A0013255/c
LOCUS      169 bp      DNA      GSS      06-JUN-1998
DEFINITION CIT-HSP-2299C22.TF CIT-HSP Homo sapiens genomic clone 2299C22,
genomic survey sequence.
ACCESSION  A0013255
VERSION    A0013255.1 GI:3185820
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 169)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
BUILDING (1998)
JOURNAL   Unpublished (1998)
COMMENT   Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..169
     other_GSS:      CIT-HSP-2299C22.TF
     Contact: Mark Adams
     Department of Eukaryotic Genomics
     The Institute for Genomic Research
     9712 Medical Center Dr., Rockville, MD 20850, USA
     Tel: 301 838 0200
     Fax: 301 838 0208
     Email: mdadams@tigr.org
     Clones are available from Research Genetics (info@resgen.com). BAC
     end search page:
     http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
     Seq primer: M13-21
     Class: BAC ends.
BASE COUNT      32 a 47 c 48 g 42 t
ORIGIN

Query Match      44.7%; Score 67; DB 96; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.2e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
DB 73 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 132
QY 111 GGGGGCC 117
DB 133 GGGGGCC 139

RESULT 10
T48593
LOCUS      259 bp      mRNA      EST      02-FEB-1995
DEFINITION Ph6f4_19/1TV Outward Alu-primed hncDNA library Homo sapiens cDNA
            clone ph6f4_19/1TV, mRNA sequence.

```

```

QY 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
DB 101 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 42
QY 111 GGGGGCC 117
DB 41 GGGGGCC 35

RESULT 9
A0041198
LOCUS      175 bp      DNA      GSS      14-JUL-1998
DEFINITION CIT-HSP-2335D21.TR CIT-HSP Homo sapiens genomic clone 2335D21,
genomic survey sequence.
ACCESSION  A0041198
VERSION    A0041198.1 GI:3310469
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 175)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
BUILDING (1998)
JOURNAL   Unpublished (1998)
COMMENT   Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..175
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="2335D21"
     /clone_lib="CIT-HSP"
     /sex="Male"
     /cell_type="Sperm"
     /note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT      43 a 50 c 47 g 35 t
ORIGIN

Query Match      44.7%; Score 67; DB 96; Length 175;
Best Local Similarity 100.0%; Pred. No. 1.2e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
DB 73 GTGGATCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 132
QY 111 GGGGGCC 117
DB 133 GGGGGCC 139

RESULT 10
T48593
LOCUS      259 bp      mRNA      EST      02-FEB-1995
DEFINITION Ph6f4_19/1TV Outward Alu-primed hncDNA library Homo sapiens cDNA
            clone ph6f4_19/1TV, mRNA sequence.

```

```

ACCESSION T48593
VERSION T48593.1 GI:642793
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 259)
Vinogradova,T.V., Lebedev,Y.B., Kopantzev,E.P., Wagner,L.L.,
Volik,S.V., Ermolaeva,O.D., Lavrentyeva,I., Monastyrskaya,G.S. and
Sverdlov,E.D.
Outward Alu-primed hncDNA library
Unpublished (1995)
Contact: Sverdlov ED
Structure and Function of Human Genes
Shemyakin Institute of Bioorganic Chemistry
16/10 Mklukho-Maklaya, Moscow, 117871, Russia
Tel: 70953306529
Fax: 70953306538
Email: sverdhungen.siocb.msk.su.
FEATURES
source
1..259
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ph6f4_19/1IV"
/clone_lib="Outward Alu-primed hncDNA library"
/notes="Vector: pGEM-3Z; Site_1: EcoRI; Site_2: BamHI; The
library was constructed as described in [Obradovic, D.,
Borodin, A.M., Kopantzev, E.P., Wagner, L.L., Volik, S.V.,
Ermolaeva, O.D., Lebedev, Y.B., Monastyrskaya, G.S.,
Sverdlov, E.D. (1993) Bioorganicheskaya khimiya, 20,
919-930]. This protocol is based on nested primer strategy
using Alu- specific primers (ALN3 and TC-65) that direct
the hncDNA synthesis outward of Alu repeats."
BASE COUNT 62 a 66 c 72 g 59 t
ORIGIN
Query Match 44.7%; Score 67; DB 20; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.3e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 GTGATCCCCGGCTCGAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 56 GTGATCCCCGGCTCGAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 115
QY 111 GGGGGCC 117
|||||
Db 116 GGGGGCC 122
RESULT 11
LOCUS AA052885 275 bp mRNA EST 13-SEP-1996
DEFINITION T3551 Bloodstream form of serodeme ILTat1.1 Trypanosoma brucei
brucei cDNA 5', mRNA sequence.
ACCESSION AA052885
VERSION AA052885.1 GI:1543898
KEYWORDS EST.
SOURCE Trypanosoma brucei brucei.
ORGANISM Trypanosoma brucei brucei.
REFERENCE Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
AUTHORS 1 (bases 1 to 275)
Osanya,A., Murphy,N.B. and Pelle,R.
TITLE Trypanosoma brucei cDNAs
JOURNAL Unpublished (1996)
COMMENT On Nov 29, 1993 this sequence version replaced gi:637578.
Contact: Osanya A
Unit 3
International Livestock Research Institute
Box 30709, Nairobi, Kenya
Tel: 254 2 630 743
Fax: 254 2 631 499
Email: a.osanya@cnet.com
Seq primer: M13 primer.
FEATURES
source
1..275
Location/Qualifiers
/organism="Trypanosoma brucei brucei"
/db_xref="taxon:5702"
/clone_lib="Bloodstream form of serodeme ILTat1.1"
/notes="cDNAs were generated from poly (A+) enriched mRNA
prepared from different developmental stages of T.b.brucei
by reverse transcription followed by PCR amplification
using mini-exon and oligo(dT) primers. The cDNA generated
were utilized in RADES-PCR coupled with differential
hybridisations to identify differentially expressed mRNA
transcripts. The products which showed to be
differentially expressed were cloned pGEM -T vector.
These differentially expressed mRNA transcripts were (are
being) sequenced to generate differentially expressed
sequence tags."
BASE COUNT 61 a 84 c 74 g 56 t
ORIGIN
Query Match 44.7%; Score 67; DB 27; Length 275;
Best Local Similarity 100.0%; Pred. No. 1.3e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 GTGATCCCCGGCTCGAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 49 GTGATCCCCGGCTCGAGGATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 108
QY 111 GGGGGCC 117
|||||
Db 109 GGGGGCC 115
RESULT 12
LOCUS B74162/c 386 bp DNA GSS 24-OCT-1998
DEFINITION CIT-HSP-2028N1.TF CIT-HSP Homo sapiens genomic clone 2028N1,
genomic survey sequence.
ACCESSION B74162
VERSION B74162.1 GI:2769849
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 386)
Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,B., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
Use of a random BAC End sequence Database for Sequence-Ready Map
Building
Unpublished (1997)
JOURNAL Other_GSSs: CIT-HSP-2028N1.TR
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21
Class: BAC ends.
FEATURES
source
1..386
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:7048251"

```



```
/db_xref="taxon:9606"
/clone="2028N1"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/notes="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT      81 a   98 c   109 g   98 t
ORIGIN

Query Match      44.7%; Score 67; DB 84; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.3e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 82 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 23
|||||
QY 111 GGGGGCC 117
|||||
Db 22 GGGGGCC 16

RESULT 13
LOCUS      AQ014512/c
DEFINITION CIT-HSP-2300D21.TF CIT-HSP Homo sapiens genomic clone 2300D21,
genomic survey sequence.
ACCESSION  AQ014512
VERSION     AQ014512.1 GI:3184839
KEYWORDS   GSS.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 462)
AUTHORS     Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE       Use of a random BAC End Sequence Database for Sequence-Ready Map
BUILDING    Building (1998)
JOURNAL     Unpublished (1998)
COMMENT     Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES             Location/Qualifiers
     source
     1..462
         /organism="Homo sapiens"
         /db_xref="GDB:7154425"
         /db_xref="taxon:9606"
         /clone="2300D21"
         /clone_lib="CIT-HSP"
         /sex="Male"
         /cell_type="Sperm"
         /notes="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT      89 a   127 c   131 g   115 t
ORIGIN

Query Match      44.7%; Score 67; DB 96; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 82 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 23
|||||
QY 111 GGGGGCC 117
|||||
Db 22 GGGGGCC 16

RESULT 14
LOCUS      B65534/c
DEFINITION CIT-HSP-2022D24.TF CIT-HSP Homo sapiens genomic clone 2022D24,
genomic survey sequence.
ACCESSION  B65534
VERSION     B65534.1 GI:2639512
KEYWORDS   GSS.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 519)
AUTHORS     Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE       Use of a random BAC End Sequence Database for Sequence-Ready Map
BUILDING    Building (1997)
JOURNAL     Unpublished (1997)
COMMENT     Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES             Location/Qualifiers
     source
     1..519
         /organism="Homo sapiens"
         /db_xref="GDB:7045730"
         /db_xref="taxon:9606"
         /clone="2022D24"
         /clone_lib="CIT-HSP"
         /sex="Male"
         /cell_type="Sperm"
         /notes="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT      110 a   137 c   148 g   124 t
ORIGIN

Query Match      44.7%; Score 67; DB 84; Length 519;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 110
|||||
Db 82 GTGGATCCCCGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGG 23
|||||
QY 111 GGGGGCC 117
|||||
Db 22 GGGGGCC 16

RESULT 15
LOCUS      AQ012629
DEFINITION AQ012629
Genomic survey sequence.
ACCESSION  AQ012629
VERSION     AQ012629.1 GI:1262912
KEYWORDS   GSS.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 547)
AUTHORS     Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE       Use of a random BAC End Sequence Database for Sequence-Ready Map
BUILDING    Building (1998)
JOURNAL     Unpublished (1998)
COMMENT     Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES             Location/Qualifiers
     source
     1..547
         /organism="Homo sapiens"
         /db_xref="GDB:7154425"
         /db_xref="taxon:9606"
         /clone="2300D21"
         /clone_lib="CIT-HSP"
         /sex="Male"
         /cell_type="Sperm"
         /notes="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT      89 a   127 c   131 g   115 t
ORIGIN
```

DEFINITION CIT-HSP-2299F1.TR CIT-HSP-Homo sapiens genomic clone 2299F1,
 genomic survey sequence.
 ACCESSION AQ012629
 VERSION AQ012629.1 GI:3185194
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 547)
 AUTHORS Adams M.D., Rounsley S.D., Zhao S., Field C.E., Bass S., Linher K.,
 Golden K., Berry K., Granger D., Suh E., Wible C., Shizuya H.,
 Simon M. and Venter J.C.
 TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
 Building (1998)
 JOURNAL Unpublished (1998)
 COMMENT Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13 Reverse
 Class: BAC ends.

FEATURES
 source
 1..547
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="GDB:7154069"
 /db_xref="taxon:9606"
 /clone="2299F1"
 /clone_lib="CIT-Hsp"
 /sex="Male"
 /cell_type="Sperm"
 /note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
 HindIII"
 BASE COUNT 154 a 139 c 125 g 129 t
 ORIGIN

Query Match 44.7%; Score 67; DB 96; Length 547;
 Best Local Similarity 100.0%; Pred. No. 1.4e-12;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 51 GTGGATCCCCGGGCTCAGGAATTCGATATCAAGCTATCGATACCGTCGACCTCGAGG 110
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 75 GTGGATCCCCGGGCTCAGGAATTCGATATCAAGCTATCGATACCGTCGACCTCGAGG 134
 QY 111 GGGGGCC 117
 |||||
 Db 135 GGGGGCC 141

Search completed: May 29, 2000, 21:13:17
 Job time: 36516 sec